## REMARKS

This Reply is responsive to the non-final Office Action dated July 15, 2005. Entry of the amendments and remarks submitted herein and reconsideration of the claimed subject matter pursuant to 37 CFR §1.112 is respectfully requested.

### I. Status of the Claims

Claims 23, 25, 26, 28-38, 40-42, 44-48 and 56-65 were pending in this application at the time of the Office Action dated July 15, 2005. Claims 35, 36, 40 and 44 were withdrawn from consideration. As a result of this amendment, new claims 66 and 67 have been added. Accordingly, claims 23, 25, 26, 28-34, 41-42, 45-48 and 56-67 are now pending and under examination.

### II. Amendments to the Claims

Claim 23 has been amended to delete clauses (a) and (e) and reference to SEQ ID NO: 9. The material of clause (e) deleted from claim 1 has been resubmitted in new claim 67. In addition, functional language similar to that seen in clause (c) of claim 23 (as amended) has been added to clause (b). Clause (b) has also been amended to indicate that the conditions recited therein are those of high stringency, with support being found in the substitute specification at page 5, lines 1-11.

Claim 26 has been amended to depend on claim 23, and to clarify that the sequence of the claimed fragment is a contiguous sequence found in any one of the recited sequences.

Applicants believe that this is clear from the disclosure at page 8, lines 17-19. Claims 28 and 61 have been amended to maintain consistency with amended claim 23. And new claim 66 has been added that indicates that the claimed nucleic acid is from *Arabidopsis*. Support for new claim 66 may be found at page 4, line 27.

No prohibited new matter has been added by way of these amendments.

## III. Claim Objections

Claim 28 has been amended to delete the word "sequence." Accordingly, the objection to claim 28 may be withdrawn.

## IV. Rejection under 35 U.S.C. §112, Second Paragraph

Claim 26 was rejected under 35 U.S.C. §112, second paragraph for alleged indefiniteness for depending on a canceled claim. Claim 26 has been amended to depend on claim 23, therefore, the rejection may be withdrawn

## V. Enablement Rejection under 35 U.S.C. §112, First Paragraph

Claims 23, 25-26, 28-34, 37-38, 41-42, 45-48 and 56-65 were rejected under 35 U.S.C. §112, first paragraph because the specification, while being enabled for an isolated nucleic acid encoding SEQ ID No. 8, the nucleic acid of SEQ ID NO: 1, host cells/plants/plant parts/seeds comprising the same and methods for transforming a plant with the same, allegedly fails to provide enablement for a nucleic acid sequence encoding any plant or animal nuclear base transporter in general or fragments thereof. In particular, the Examiner maintains the position

that Applicant does not provide any guidance for the obtention and use of nucleic acids across the entire scope claimed, and that the specification is silent regarding the functional activity of the other sequences disclosed in the specification. The Examiner continues to assert that no guidance has been provided for regions in SEQ ID No. 1 that would tolerate modifications. Further, the Examiner asserts that no specific methods are known or disclosed in the specification which provide nucleic acids of claim 23, parts (a), (c), and (d), or claims 26 and 45-46. Applicants respectfully traverse the rejection.

First, Applicants respectfully note that claim 23 has been amended to delete reference to animal nuclear base transporters, and reference to nucleic acids encoding nuclear base transporters that may be isolated via complementation in yeast. Accordingly, claim 23 as amended is now directed to an isolated nucleic acid that codes for a plant nuclear base transporter comprising a nucleic acid with a sequence that codes for a protein having a sequence according to SEQ ID NO: 8, and nucleic acids which code for a polypeptide or protein with nuclear base transporter activity that hybridize with a nucleic acid that codes for a protein having a sequence according to SEQ ID NO: 8 and which code for a polypeptide or protein with nuclear base transporter activity having at least 40% identity to SEQ ID NO: 8. Thus, the claims are directed to one of the plant nuclear base transporter nucleic acid sequences disclosed in the application, and plant nuclear base transporter sequences that hybridize to that sequence or encode nuclear base transporter proteins having a high level of identity with that sequence.

Notwithstanding the claim amendments submitted above, the Office Action is incorrect in the assertion that no specific methods are known or disclosed in the specification which provide nucleic acids of claim 23, parts (a), (c), and (d), or claims 26 and 45-46. For instance, claim 23,

part (a) (now canceled), was directed to a nucleic acid that is obtainable through complementation of nuclear base transporter-deficient yeast cells with a plant or animal gene bank and selection of nuclear base transporter-positive host cells. As described in Example 1, the PUP1 nuclear base transporter gene (that encodes the protein having SEQ ID NO: 8) was cloned by complementation of a transporter deficient yeast cell with a cDNA gene bank from *Arabidopsis thaliana* with selection for nuclear base transporter-positive cells. *Arabidopsis thaliana* is commonly used as an exemplary plant species due to its rapid development and small growing space. Thus, Applicants clearly disclose a plant nucleic acid that is obtainable through complementation of nuclear base transporter-deficient yeast cells as recited in original claim 23, part (a).

Prior to the amendment above, claim 23, part (c) was directed to a nucleic acid that hybridizes with a nucleic acid according to part (a) in a solution comprising 25% formamide, 5X SSPE, 0.1% SDS, 5X Denhardt and 50 μg herring-sperm DNA after 20 hours at 37°C, followed by washing in 2X SSC and 0.1% SDS at 42°C. Given the disclosed sequence of the PUP1 gene (SEQ ID NO: 1), the skilled artisan could clearly design a sequence that hybridizes under the recited conditions to the disclosed sequence absent undue experimentation.

Further, prior to the amendment, claim 23, part (d) was directed to a nucleic acid which codes for a polypeptide or protein with nuclear base transporter activity having at least 40% identity to SEQ ID NO: 8. As noted in the previous Reply, the other nucleic acids disclosed in the sequence listing and recited in the claims were identified by homology analysis and would predictably have properties similar to the nucleic acid having the sequence SEQ ID NO: 1 (which encodes a protein having the amino acid sequence of SEQ ID NO: 8). These sequences

have at least 40% identity to SEQ ID NO: 8, and in fact it has been confirmed that at least one of these sequences has nuclear base transporter activity. For instance, as reported in the attached publication by Bürkle et al., the PUP2 gene (SEQ ID NO: 2 of the present application) is shown to encode a protein that mediates cytokinin transport (see abstract) and has 58% identity to PUP1 at the amino acid level (see p. 17, col. 1). Thus, Applicants disclose at least two and potentially eight nucleic acids which code for a polypeptide or protein with nuclear base transporter activity having at least 40% identity to SEQ ID NO: 8.

Claim 26 is directed to a fragment of the nucleic acid of Claim 23, wherein said fragment inhibits the expression of a nuclear base transporter in a host cell when expressed from a promoter in the antisense orientation, and wherein said fragment is at least ten nucleotides, and wherein said fragment consists of a nucleic acid sequence contained within a sequence selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, and 10. Claims 45 and 46 specify that the claimed fragment is at least 50 nucleotides, or at least 200 nucleotides, respectively. As discussed above, at least two of the disclosed sequences have been demonstrated to code for proteins having nuclear base transporter activity. The other sequences encode proteins having at least 40% identity, and are predicted to encode proteins having a similar function. It would be well within the ability of one of ordinary skill in the art to design fragments of these disclosed sequences having the recited lengths that inhibit protein activity when expressed in the antisense orientation.

According to the Office Action, Newman et al. is relied upon to provide evidence that not every 10, 50 and 200 nucleotides of SEQ ID NO: 1 is enabled or has function, since Newman et al. disclose a nucleic acid with more than 160 contiguous bases of SEQ ID NO: 1 having no

known base transport activity. Applicants respectfully note that claims 26, 45 and 46 are not directed to fragments having base transport activity but to fragments that inhibit base transport activity when expressed in the antisense orientation. Given the disclosed sequences and the disclosed function of the encoded proteins, the skilled artisan could clearly select a fragment of 10, 50 or 200 nucleotides that inhibits the base transport of the disclosed proteins when expressed in the antisense orientation.

The Office Action further asserts that Applicants cite no reference or scientific publication that teaches multiple site mutagenesis in a specific gene that retains the original gene function, and that Applicant points to no suitable yeast systems for the selection and identification of the nucleic acids as broadly as claimed. This is incorrect, since Applicants clearly disclose a nuclear base transporter deficient yeast strain that may be used to screen for functional nuclear base transporter genes as described in Example 1 of the specification. Further, methods of mutagenesis whereby a given gene may be modified to contain random, multiple mutations have been standard practice in the art for years as evidenced by several attached publications.

For instance, as reported by Pjura et al. (Protein Science, 1993, 2: 2217-2225), eighteen functional, thermostable single site T4 lysozyme mutants were isolated using a variety of random mutagenesis techniques, including hydroxylamine mutagenesis, thionucleotide misincorporation mutageneis and avian myeloma virus (AMV) reverse transcriptase misincorporation mutagenesis (see p. 2218). According to the authors, AMV reverse transcriptase mutagenesis is prone to making more than one mutation in a gene (p. 2221, col. 2).

In the article by Uppaluri and Towle (Mol. Cell. Biol., 1995, 1499-1512), the authors screened random mutations generated by mutagenic PCR in the rat hormone receptor β, and identified a group of "superactivator" mutations leading to enhanced hormone-dependent reporter gene activation in the yeast *S. cerevisiae* (see the abstract and p. 1500, col. 2). As shown in Table 3 on page 1505, several of these mutants have more than one mutation.

In the article by Kim et al. (J. Biol. Chem., 1996, 271(9): 4872-78), the authors report the isolation of catalytically active mutants of HIV reverse transcriptase generated by random sequence mutagenesis coupled with positive selection in *E. coli* (see abstract). The positive selection screen was conducted by selecting functional mutants that were able to complement the temperature sensitive phenotype of an *E. coli* polymerase mutant (see p. 4874). This reference shows that it is clearly within the ordinary skill of the artisan to isolate random, active mutants by functional complementation in another species.

Given these well known mutagenesis methods and the screening assay disclosed in the specification, one of ordinary skill could readily isolate libraries of variants of the disclosed sequences containing one or more mutations, and screen these libraries for variants demonstrating functional nuclear base transporter activity as recited in the claims.

In summary, as noted above, claim 23 has been amended to delete reference to human nuclear base transporter genes and nucleic acids that are obtained through complementation of yeast nuclear base transporter-deficient host cells. Since the application clearly discloses at least two and potentially eight nucleic acids encoding polypeptides having nuclear base transporter activity with at least 40% identity to SEQ ID NO: 8 as discussed above, and the skilled artisan could certainly design and isolate functional variants, inhibitory complementary fragments, as

well as nucleic acids that hybridize to a sequence encoding a protein having SEQ ID NO: 8 under the recited conditions absent undue experimentation, the enablement rejection under §112 should now be withdrawn.

## VI. Written Description Rejection under 35 U.S.C. §112

Claims 23, 25-26, 28-34, 37-38, 41-42, 45-48 and 56-65 were rejected under 35 U.S.C. §112, first paragraph, as containing subject matter which allegedly was not described in the specification in such a way as to reasonably convey to the skilled artisan that the inventors had possession of the claimed invention at the time the application was filed. Essentially, the Examiner criticizes the description of the nucleic acids of claim 23, part (a) as being described by their activity only. The Examiner also asserts that a substantial variation in structures and function would be expected among the hybridizing sequences of part (c) because the hybridization conditions recited allegedly define low stringency. The Examiner further asserts that structural differences would be expected among nucleic acids of part (d) encoding proteins with as low as 40% sequence identity to SEQ ID NO: 8. Applicants respectfully traverse the rejection.

First, Applicants respectfully note that part (a) of claim 23 has been deleted by way of amendment above. Therefore, the Examiner's concerns regarding the scope of this part of the claim are now moot.

Second, Applicants respectfully disagree that the hybridization conditions recited in claim 23 are those of low stringency. For instance, as disclosed on page 5, lines 1-11 of the substitute specification, the recited conditions exemplify extremely stringent hybridization

conditions. The Examiner has provided no reasoning or evidence to challenge the assertion that the recited conditions result in stringent hybridization. Claim 23 has been amended to clarify that the recited conditions are those of high stringency.

Third, although the Examiner is correct to note that structural differences would be expected among nucleic acids of part (d) encoding proteins with at least 40% sequence identity to SEQ ID NO: 8, part (d) (now part (c) following the amendments made above) also recites functional language in that the claimed nucleic acid codes for a polypeptide or protein with nuclear base transporter activity. Applicants have demonstrated that nucleic acids showing such nuclear base transporter activity with such levels of identity to SEQ ID NO: 8 can be isolated by homology analysis. For instance, as reported in the attached publication by Bürkle et al., the PUP2 gene (SEQ ID NO: 2 of the present application) is shown to encode a protein that also mediates cytokinin transport (see abstract) and has 58% identity to PUP1 (SEQ ID NO: 8) at the amino acid level (see p. 17, col. 1).

In view of the amendments and remarks made above, Applicants respectfully request reconsideration and withdrawal of the written description rejection under §112, first paragraph.

## VII. Prior Art Rejections

Claims 23, 25, 26, 45-46, 57 and 62-63 were rejected under 35 U.S.C. §102(b) as being allegedly anticipated by, or in the alternative, under 35 U.S.C. §103(a) as obvious over Schultes et al. According to the Office Action, Schultes et al. teach an isolated gene from maize with similarity to pyrimidine and purine transport proteins. Applicants respectfully traverse the rejection with respect to the amended claims above.

As discussed above, claim 23 as amended is now directed to an isolated nucleic acid that codes for a plant nuclear base transporter comprising a nucleic acid with a sequence that codes for a protein having a sequence according to SEQ ID NO: 8, and nucleic acids which code for a polypeptide or protein with nuclear base transporter activity that hybridize with a nucleic acid that codes for a protein having a sequence according to SEQ ID NO: 8 and which code for a polypeptide or protein with nuclear base transporter activity having at least 40% identity to SEQ ID NO: 8. Thus, the claims are directed to one of the specific plant (*Arabidopsis*) nuclear base transporter nucleic acid sequences disclosed in the application, and plant nuclear base transporter sequences that hybridize to that sequence or encode nuclear base transporter proteins having a high level of identity with that sequence.

The protein disclosed in Schultes et al. does not appear to meet the limitations of amended claim 23. In addition, claim 26 has been amended to indicate that the claimed fragment is at least ten contiguous nucleotides from any one of the sequences selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, and 10. In view of these amendments, reconsideration and withdrawal of the rejection under 35 U.S.C. §102(b)/§103(a) based on Schultes et al. are respectfully requested.

Claims 23, 25, 26, 28, 30-32, 41, 45-46, 48, 57 and 62 were rejected under 35 U.S.C. §102(e) as being allegedly anticipated by, or in the alternative, under 35 U.S.C. §103(a) as obvious over Abramson et al. According to the Office Action, Abramson et al. teach an isolated human or rat urate transporter. Applicants respectfully traverse the rejection with respect to the amended claims above.

As discussed above, claim 23 as amended is now directed to an isolated nucleic acid that codes for a plant nuclear base transporter. As animal nuclear base transporter nucleic acids are no longer included in the amended claims, withdrawal of the rejection under 35 U.S.C. §102(b)/§103(a) based on Abramson et al. is respectfully requested.

This reply is fully responsive to the Office Action dated July 15, 2005. Therefore, a Notice of Allowance is next in order and is respectfully requested.

Except for issue fees payable under 37 CFR §1.18, the commissioner is hereby authorized by this paper to charge any additional fees during the pendency of this application including fees due under 37 CFR §1.16 and 1.17 which may be required, including any required extension of time fees, or credit any overpayment to Deposit Account 50-0310. This paragraph is intended to be a **CONSTRUCTIVE PETITION FOR EXTENSION OF TIME** in accordance with 37 CFR §1.136(a)(3).

If the Examiner has any further questions relating to this Reply or to the application in general, she is respectfully requested to contact the undersigned by telephone so that allowance of the present application may be expedited.

Respectfully submitted,

**MORGAN, LEWIS & BOCKIUS LLP**

Bonnie Weiss McLeod
Reg. No. 43,255

Dated: October 17, 2005

**Customer No. 009629**
**MORGAN, LEWIS & BOCKIUS LLP**
1111 Pennsylvania Avenue, NW
Washington, D.C. 20004
(202) 739-3000

*The Plant Journal* (2003) **34**, 13–26

# Transport of cytokinins mediated by purine transporters of the PUP family expressed in phloem, hydathodes, and pollen of *Arabidopsis*

**Lukas Bürkle[†,§], Anna Cedzich[§], Corinna Döpke, Harald Stransky, Sakiko Okumoto, Bernd Gillissen[‡], Christina Kühn and Wolf B. Frommer[1,*]**

*Plant Physiology, Zentrum für Molekularbiologie der Pflanzen (ZMBP), Auf der Morgenstelle 1, D-72076 Tübingen, Germany, and*

*[1]Carnegie Institution of Washington, 260 Panama Street, Stanford, CA 94305, USA*

Received 21 November 2002; accepted 19 December 2002.

*For correspondence (fax +49 7071 29 3287; e-mail frommer@andrew2.stanford.edu).

[†]Present address: Max-Planck-Institut für Molekulare Genetik, Ihnestr. 73, D-14195 Berlin, Germany.

[‡]Present address: Robert-Rössle-Klinik, Lindenberger Weg 80, D-13125 Berlin, Germany.

[§]These authors contributed equally to this study.

## Summary

**Nucleobases and derivatives like cytokinins and caffeine are translocated in the plant vascular system. Transport studies in cultured *Arabidopsis* cells indicate that adenine and cytokinin are transported by a common $H^+$-coupled high-affinity purine transport system. Transport properties are similar to that of *Arabidopsis* purine transporters AtPUP1 and 2. When expressed in yeast, AtPUP1 and 2 mediate energy-dependent high-affinity adenine uptake, whereas AtPUP3 activity was not detectable. Similar to the results from cell cultures, *pu*rine *p*ermeases (PUP) mediated uptake of adenine can be inhibited by cytokinins, indicating that cytokinins are transport substrates. Direct measurements demonstrate that AtPUP1 is capable of mediating uptake of radiolabeled *trans*-zeatin. Cytokinin uptake is strongly inhibited by adenine and isopentenyladenine but is poorly inhibited by 6-chloropurine. A number of physiological cytokinins including *trans*- and *cis*-zeatin are also efficient competitors for AtPUP2-mediated adenine uptake, suggesting that AtPUP2 is also able to mediate cytokinin transport. Furthermore, AtPUP1 mediates transport of caffeine and ribosylated purine derivatives in yeast. Promoter–reporter gene studies point towards AtPUP1 expression in the epithem of hydathodes and the stigma surface of siliques, suggesting a role in retrieval of cytokinins from xylem sap to prevent loss during guttation. The *AtPUP2* promoter drives GUS reporter gene activity in the phloem of *Arabidopsis* leaves, indicating a role in long-distance transport of adenine and cytokinins. Promoter activity of *AtPUP3* was only found in pollen. In summary, three closely related PUPs are differentially expressed in *Arabidopsis* and at least two PUPs have properties similar to the adenine and cytokinin transport system identified in *Arabidopsis* cell cultures.**

**Keywords: nucleobase, adenine, caffeine, adenosine, hormone, transport.**

## Introduction

As a result of their immobile nature, plants require highly efficient mechanisms for adaptation to rapidly changing environmental conditions and for communication between the distal organs of the plant. Long-distance signaling is preferentially mediated by chemical signals as opposed to electrical signaling used by animals. It is therefore not surprising that besides classical hormones, such as steroids, oligopeptides, and eicosanoid-like compounds, plants have developed a specific set of phytohormones. Most phytohormones are synthesized by a few conversions from normal metabolic intermediates. The phytohormone cytokinin is generated by prenylation of nucleosides or nucleotides by cytokinin synthase (Astot *et al.*, 2000; Kakimoto, 2001; Takei *et al.*, 2001). Cytokinins not only control cell division, but affect many physiological and developmental functions, e.g. leaf senescence, nutrient mobilization and biomass distribution, apical dominance, formation and activity of shoot apical meristems, floral transition,

breaking of bud dormancy, and seed germination (Mok and Mok, 2001). Root apical meristems seem to represent major sites of the synthesis of free cytokinins, however, young leaves and embryos can also serve as sources. Cytokinins are transported from roots to shoots via the xylem, additionally reflux occurs in the phloem, supporting the hypothesis that they serve as mobile signaling molecules (Beveridge *et al.*, 1997; Emery *et al.*, 2000; Weiler and Ziegler, 1981). Thus, it is obvious that multiple cellular importers and exporters are required to allow efficient mobilization and targeted translocation. However, in contrast to the mechanisms of polar transport of the hormone auxin, basically nothing is known about the mechanisms of cytokinin transport (Swarup *et al.*, 2000). When cell cultures are exposed to free cytokinin bases, these are rapidly taken up and inactivated by glycosylation and stored in the vacuole (Fußeder *et al.*, 1989). *Cis*- and *trans*-zeatin *O*-glucosyl and *O*-xylosyl transferase genes have been identified, and will help towards a better understanding of phytohormone metabolism (Martin *et al.*, 1999a,b, 2001).

At elevated atmospheric $CO_2$, plants undergo an accelerated life cycle leading to increased nitrogen requirements, uptake and mobilization, and to increased leaf expansion (Ludewig and Sonnewald, 2000). One cause of the higher growth rates could be increased production and transport of cytokinins as compared to ambient conditions. In fact, increased concentrations of cytokinins were found in xylem exudate of plants grown in elevated $CO_2$ (Yong *et al.*, 2000). The lack of a direct correlation between the increases of cytokinin concentration in xylem sap and leaf contents indicates a complex interaction between transport and metabolism along the translocation path.

Recently, the enzymes responsible for cytokinin metabolism, i.e. cytokinin oxidase and corresponding genes, have been identified (Bilyeu *et al.*, 2001; Houba-Hérin *et al.*, 1999). Ectopic overexpression of several *Arabidopsis* cytokinin oxidase genes had a dramatic effect on the shoot to root ratios and on the leaf cell number (Werner *et al.*, 2001). The presence of signal sequences together with yeast expression data may indicate that some of the enzymes are secreted. Besides a role in cytokinin degradation, cytokinin oxidases thus may also control cytokinin concentrations available to the plasma membrane cytokinin receptors. Candidates for plasma membrane receptors belonging to the class of two component histidine kinase systems (Inoue *et al.*, 2001; Kakimoto, 1996; Suzuki *et al.*, 2001; Ueguchi *et al.*, 2001) have been identified. On the other hand, the presence of potential intracellular cytokinin receptors may suggest that besides extracellular signal perception, the uptake of cytokinins into cells may also be important (Brault *et al.*, 1997; Kulaeva *et al.*, 1995, 1998). Furthermore, if a significant portion of cytokinin is catabolized extracellularly, transporters are required for the recycling of adenine. Recently, nuclear response regulators and transcriptional repressors have also been identified, providing a deep insight into the cytokinin-mediated signal transduction cascade (Hwang and Sheen, 2001).

In agreement with the above findings, transport seems to be highly controlled, as a local supply of cytokinins by microorganisms on senescing leaves leads to a local delay of senescence. Furthermore, the grafting of wild-type shoots onto root stocks of transgenic tobacco overproducing cytokinins as a result of ectopic expression of the *Agrobacterium tumefaciens ipt* gene, did not lead to the release of lateral shoot buds from suppression or to accelerated senescence (Faiss *et al.*, 1997). Also, localized induction of *ipt* gene expression had only local effects, leading to a paracrine hypothesis of cytokinin action. Thus, the major step, which is least understood and still undefined at the molecular level, is the role of cytokinin and cytokinin catabolite transporters. Surprisingly, no reference was found concerning uptake measurements using radiolabeled cytokinins aiming at a characterization of plant cytokinin transport.

To identify potential cytokinin transporters, an indirect approach was chosen based on the structural similarity of cytokinins and purine bases and the necessity of the coexistence of adenine and cytokinin transport systems, if metabolism can occur in the apoplasm. A purine transport-deficient yeast mutant *fcy2* was used for suppression cloning of plant transporter genes, enabling the identification and characterization of a new superfamily of small hydrophobic polytopic membrane proteins *pu*rine *p*ermeases (PUP) mediating high-affinity transport of nucleobases (Gillissen *et al.*, 2000). Competition studies showed that AtPUP1 was also able to recognize cytokinins.

The present study describes an adenine uptake system found in cultured *Arabidopsis* cells that is inhibited by cytokinins and thus shares properties similar to that of AtPUP1. The ability of AtPUP1 to transport cytokinins was proven by direct uptake studies using radiolabeled *trans*-zeatin in yeast expressing AtPUP1. Analysis of expression using promoter–reporter fusion indicates that AtPUP1 may play a role in the retrieval of nucleobase derivatives including xylem-delivered cytokinins in the epithem of hydathodes and at the stigma surface of siliques. Two homologs of AtPUP1 were cloned and expressed in yeast. AtPUP2 mediated adenine uptake and recognized *trans*- and *cis*-zeatin, isopentenyladenine, kinetin, and benzylaminopurine as substrates. In contrast, AtPUP3 did not show detectable activity in yeast. The AtPUP2 promoter drives expression of the GUS reporter gene in the phloem of *Arabidopsis* leaves, suggesting a role in phloem loading and transport of adenine and cytokinins, whereas AtPUP3 promoter activity is restricted to pollen.

## Results

### *Uptake of adenine by* Arabidopsis *cell cultures*

To determine the properties of adenine uptake by *Arabidopsis* cell cultures, $^{14}$C-adenine transport was measured. Uptake was linear for at least 3 min and sensitive to the protonophore CCCP, suggesting a secondary active uptake mechanism (Figure 1a). Competition with a 10-fold excess of unlabeled adenine, isopentenyladenine, and *trans*-zeatin strongly reduced the uptake rate, indicating that adenine and cytokinins are taken up by a common transport system (Figure 1b). *Trans*-zeatin riboside had almost no inhibitory effect, similar as for AtPUP1 (Gillissen *et al.*, 2000). Expression of the adenine transporter AtPUP1 in the cell culture was confirmed by RT-PCR (Figure 1c).

### *Transport of cytokinins by AtPUP1*

Energy dependence and specificity indicated that a purine transporter of the PUP family may be responsible for the uptake systems found in cell cultures. Additionally, competition studies in yeast have indicated that AtPUP1 is able to recognize cytokinins. However, competition does not prove that a substance is actually transported (Gillissen *et al.*, 2000). To increase the sensitivity of the uptake system and in order to prove cytokinin transport directly, AtPUP1 was expressed in yeast under the control of a modified strong yeast $H^+$-ATPase *PMA1* promoter. For this purpose, the cDNA encoding *AtPUP1* was subcloned into pDR195 (Rentsch *et al.*, 1995). Expression of AtPUP1 in the yeast mutant MG887-1 deficient for adenine uptake revealed significantly better growth on a medium containing adenine as the sole N-source as compared to expression under control of the *PGK* promoter in the original vector pFL61 (Figure 2). This result indicates that a stronger expression leads to an increased number of AtPUP1 molecules at the plasma membrane and thus to higher transport rates for adenine.

The ability of AtPUP1 to transport cytokinin was analyzed by direct uptake measurements of radiolabeled *trans*-zeatin into strain MGG887-1 transformed with AtPUP1. A typical time course of *trans*-$^3$H-zeatin uptake is shown in Figure 3(a). The results demonstrate an uptake of *trans*-zeatin by AtPUP1 at a concentration of 100 μM, however, uptake was linear for only a short time (approximately 60 sec). In contrast, almost no uptake was detectable in yeast transformed with the empty pDR195 vector. The tritium label of radiolabeled *trans*-zeatin is located on the purine ring of *trans*-zeatin (Figure S1a). To exclude the possibility that uptake is caused by radiochemical impurities like $^3$H-adenine, HPLC analysis was performed, revealing that more than 98% of detected radioactivity had the same retention time as *trans*-zeatin (Figure S1b), indicating that the measured uptake

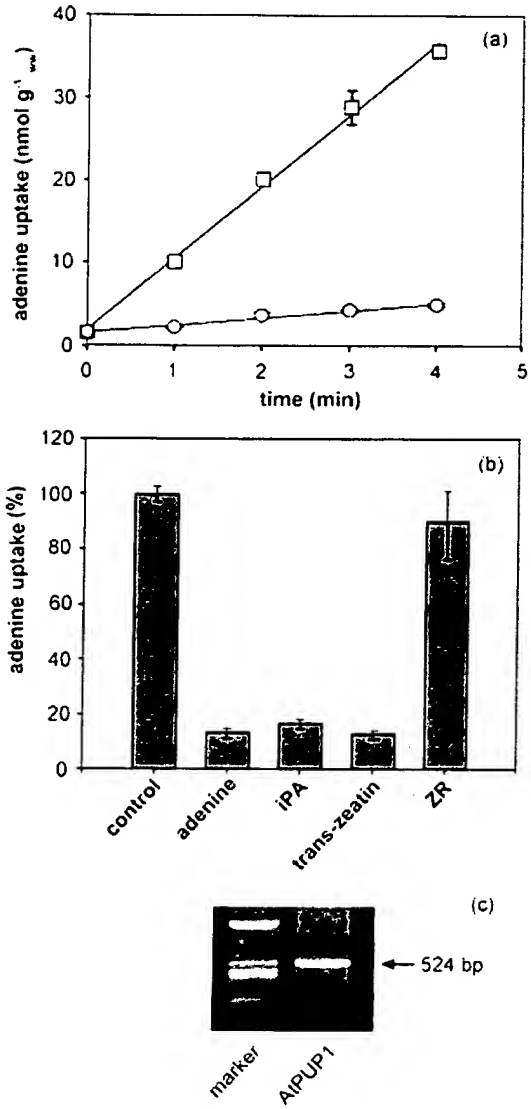



**Figure 1.** Uptake of adenine by *Arabidopsis* suspension cell culture.
(a) Time and energy dependence. *Arabidopsis* suspension cells were assayed for $^{14}$C-adenine uptake at 20 μM substrate concentration and pH 5.7 in the presence (○) or absence (□) of 100 μM carbonyl cyanide *m*-chlorophenyl-hydrazone (CCCP).
(b) Competition studies with adenine and cytokinins. Uptake of $^{14}$C-adenine was determined at 20 μM adenine in the presence of a 10-fold excess (200 μM) of unlabeled competitors. Values represent the mean ± SD of three independent experiments; WW, wet weight; iPA, isopentenyladenine; ZR, *trans*-zeatin riboside.
(c) Analysis of AtPUP1 expression by RT-PCR. RNA from suspension cells was converted to cDNA by reverse transcription. A 524 bp fragment of *AtPUP1* was amplified by 30 PCR cycles.

corresponds to transport of *trans*-zeatin by AtPUP1. Using only the initial linear phase, a crude determination of the affinity indicates a $K_m$ value of approximately 40 μM similar to the $K_i$ of *trans*-zeatin and kinetin for adenine uptake by AtPUP1 determined previously (Gillissen *et al.*, 2000). To investigate the substrate specificity of AtPUP1, competition

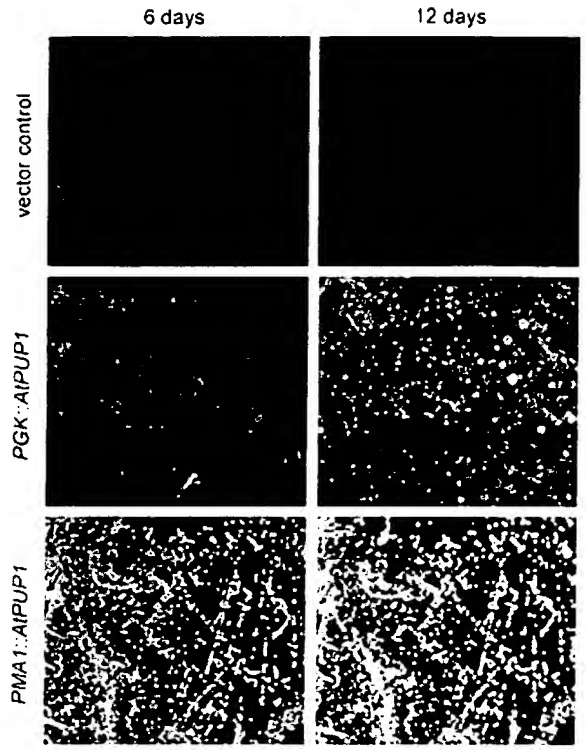



**Figure 2.** Functional complementation of the yeast adenine transport mutant MG887-1 by AtPUP1 under control of two different promoters.
MG887-1 yeast was transformed with AtPUP1 expressed under control of a *PMA1* promoter fragment in pDR195, from the *PGK* promoter in pFL61 or with the empty pDR195 vector as control. Yeast strains were plated on minimal medium containing adenine as the sole nitrogen source (1 mg $ml^{-1}$) and grown for 6 (left panel) and 12 (right panel) days.

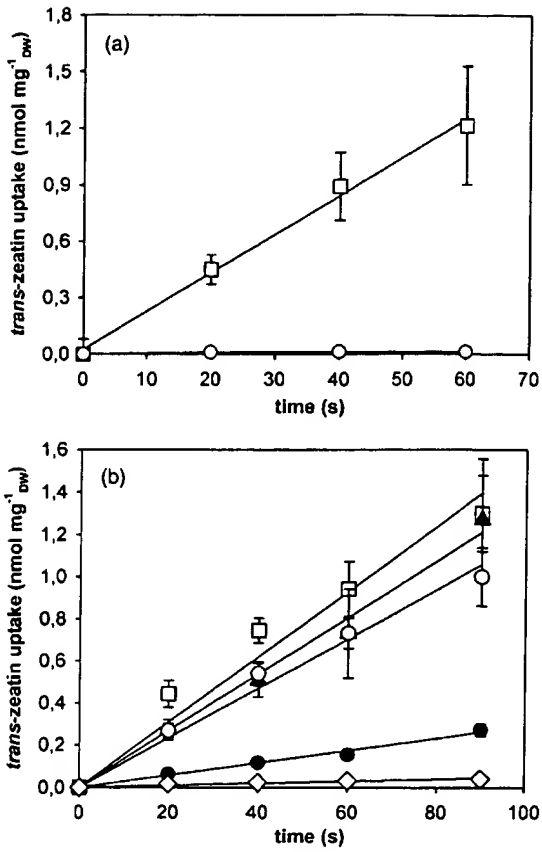



**Figure 3.** AtPUP1-mediated uptake of *trans*-zeatin in yeast.
(a) Time dependence. The uptake of *trans*-zeatin (nmol $mg^{-1}$ dry weight (DW)) was determined radioactively (*trans*-$^3$H-zeatin, specific activity: 87 Bq $\mu l^{-1}$) in the MG887-1 yeast strain transformed with AtPUP1 in pDR195 (□) or with empty vector pDR195 (○) at 100 μM *trans*-zeatin and pH 5.0.
(b) Substrate specificity of AtPUP1. Uptake of *trans*-zeatin by AtPUP1 in pDR195 was determined by competition with a 10-fold excess (1 mM) of allantoin (▲), 6-chloropurine (○), kinetin (●), and isopentenyladenine (◇), respectively. *Trans*-zeatin uptake without competitor is shown by (□). Background uptake rates in MG887-1 transformed with the empty pDR195 vectors were subtracted. Results represent the mean ± SD of three independent experiments.

experiments with other cytokinins like isopentenyladenine and kinetin, and other different structural and metabolism-related substances like 6-chloropurine, adenine, allantoin, and sucrose as a control were performed. The uptake of *trans*-zeatin was strongly inhibited by isopentenyladenine and kinetin (Figure 3b) and, as expected, also by adenine (data not shown). In contrast, sucrose had no effect on the uptake of *trans*-zeatin (data not shown), whereas allantoin and 6-chloropurine had only weak inhibitory effects. The results indicate that AtPUP1 can transport a variety of physiologically relevant cytokinins. Lack of allantoin transport activity is consistent with the finding that allantoin uptake is mediated by ureide permeases (UPS) heterocycle transporters (Desimone *et al.*, 2002).

### *Transport of caffeine by AtPUP1*

It has been suggested that caffeine has a cytokinin-like activity in different plant tissues (Vitória and Mazzafera, 1997). Caffeine inhibited adenine uptake by AtPUP1 (Gillissen *et al.*, 2000). Caffeine has inhibitory and toxic effects on the growth of *Saccharomyces cerevisiae* (Bard *et al.*, 1980; Hannan and Nasim, 1977). To test whether AtPUP1 is able to mediate the transport of caffeine, the growth inhibitory effect of caffeine on the growth of MG887-1 transformed with AtPUP1 was analyzed in the presence of ammonium as a nitrogen source. Yeast expressing AtPUP1 was significantly more sensitive to caffeine at concentrations of 0.1 and 0.2%, resulting in decreased colony size as compared to yeast transformed with the empty vector (Figure 4). The observed inhibition of growth is most probably caused by higher caffeine uptake, indicating the ability of AtPUP1 to recognize caffeine as a substrate, which is transported into the yeast cell.

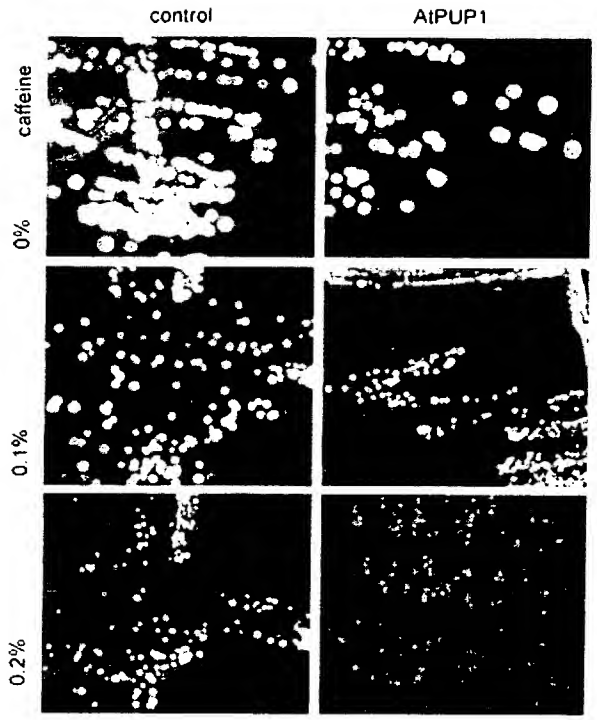



**Figure 4.** AtPUP1 confers hypersensitivity to caffeine.
MG887-1 yeast strains expressing AtPUP1 (right panel) and the vector control (left panel) were grown for 6 days on minimal medium containing caffeine at concentrations of 0, 0.1, and 0.2% (magnification in all panels 4×).

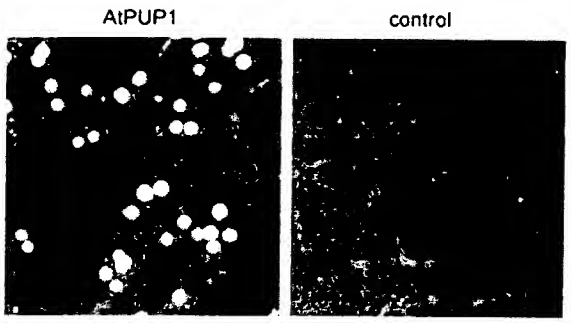



**Figure 5.** Functional complementation by AtPUP1 of the adenine auxotrophic DM734–284D yeast strain deficient in adenosine uptake.
Yeast strains expressing AtPUP1 (left panel) or containing pDR195 vector (right panel) were plated on minimal medium supplemented with 150 μM adenosine as the sole purine source and grown for 3 days.

### *Transport of nucleosides by AtPUP1*

As adenosine and ribosides of cytokinin are also transported in plants, it was investigated whether AtPUP1 transports adenosine. As *Saccharomyces cerevisiae* does not possess an efficient adenosine transport system, adenine auxotrophic strains cannot grow on adenosine as the sole purine source (Detke, 1998; Mäser *et al.*, 1999). In contrast to controls, the DM734-238D yeast strain expressing AtPUP1 in pDR195 was able to grow on adenosine-containing media, proving that AtPUP1 can transport ribosides such as adenosine (Figure 5). This finding is in agreement with the weak inhibitory effect of adenosine on adenine uptake (Gillissen *et al.*, 2000).

### *Isolation of AtPUP2 and 3 cDNAs*

Database searches and phylogenetic analyses were used to classify 20 paralogs from the completed *Arabidopsis* genome (Figure S2). Within the PUP family, AtPUP2 and 3 are the closest relatives of AtPUP1. These two proteins show 64 and 58% identity to AtPUP1 at the amino acid level, respectively (Figure S3). AtPUP1 and AtPUP3 are located on chromosome I, whereas *AtPUP2* maps to chromosome II. To study the function of these new proteins, full-length cDNAs were amplified by RT-PCR. *AtPUP2* cDNA contains an open reading frame of 1041 bp encoding a protein of 347 amino acids with a calculated molecular mass of 38.1 kDa. The cDNA of *AtPUP3* contains an open reading frame of 1053 bp encoding a protein with 351 amino acids with a calculated molecular mass of 38.9 kDa. The hydropathy patterns generated by THMM1.0 (Sonnhammer *et al.*, 1998) suggest the presence of 10 transmembrane-spanning domains with N- and C-termini being located in the cytosol (Figure S4).

### *Functional analysis of AtPUP2 and 3 in yeast*

To analyze the functionality of AtPUP2 and 3 in yeast, both cDNAs were cloned into the yeast expression vector pDR195. The capacity to transport adenine was tested by direct uptake measurements in the purine uptake-deficient yeast strain MG887-1 transformed with AtPUP2 or AtPUP3 in pDR195. AtPUP2-mediated linear uptake for $^{14}$C-adenine (Figure 6a). Uptake was low, but significantly higher when compared to controls transformed with the empty vector, demonstrating that AtPUP2 is a functional adenine transporter. However, transport rates for adenine mediated by AtPUP2 were significantly lower compared to AtPUP1, potentially because of less efficient targeting of AtPUP2 to the plasma membrane. Adenine uptake was pH dependent (Figure 6b) with highest uptake rates at pH 3.0. Using the Michaelis–Menten equation and non-linear regression analysis, a $K_m$ value of 22.6 ± 1.0 μM was determined at pH 3.5 (Figure 6c). AtPUP2 was not able to complement the growth deficiency of MG887-1 when tested on media containing adenine as the sole nitrogen source (data not shown). This is not surprising as the assay is not very sensitive. Adenine is a poor nitrogen source, therefore metabolization of adenine is rate limiting. The ability of AtPUP2 to recognize cytokinins as a substrate was tested by competition of $^{14}$C-adenine uptake in the presence of an excess of unlabeled cytokinins (Figure 6d). $^{14}$C-adenine

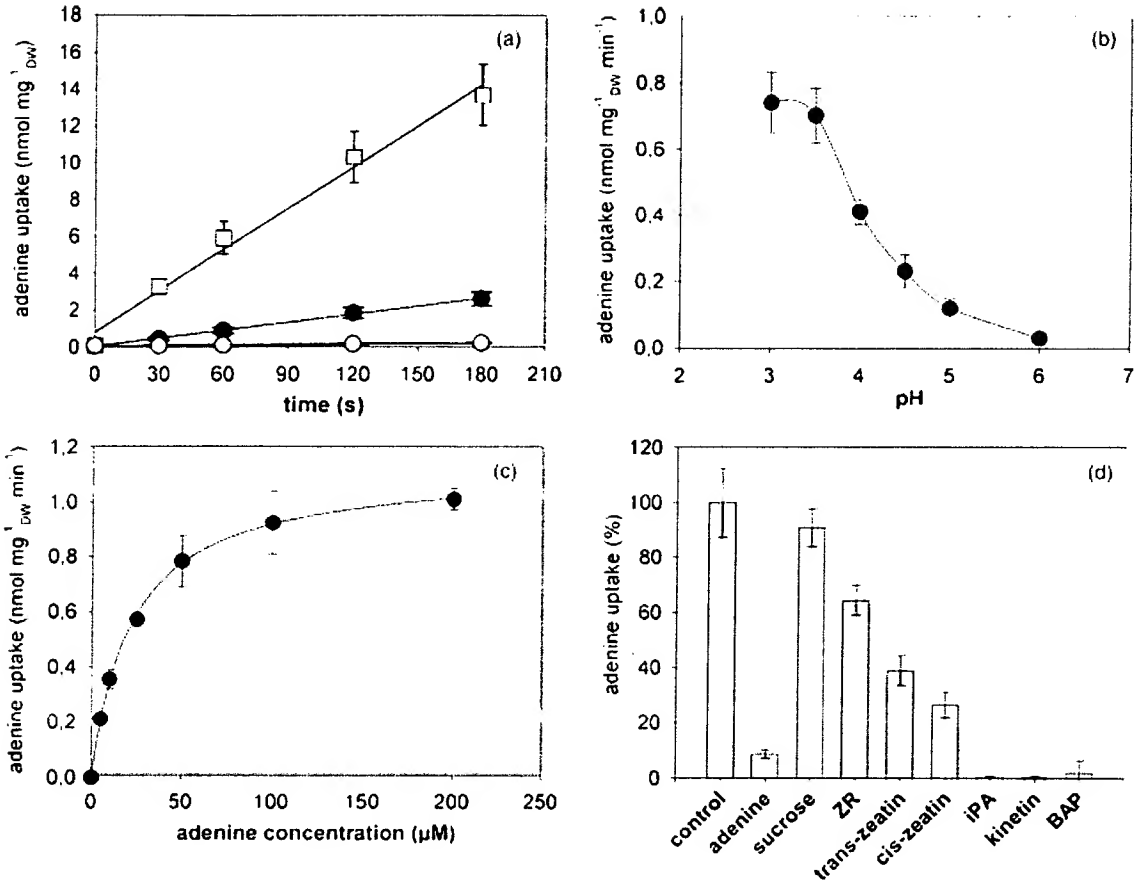



**Figure 6.** Kinetics of AtPUP2-mediated adenine uptake.
(a) Time course of adenine uptake. MG887-1/ura3 yeast transformed with AtPUP1 in pDR195 (□), AtPUP2 in pDR195 (●) or empty pDR195vector (○) were assayed for $^{14}$C-adenine uptake at 100 μM substrate concentration at pH 3.5. DW, dry weight.
(b) pH dependence of adenine uptake. Adenine uptake rates of MG887-1/ura3 yeast expressing AtPUP2 in pDR195 measured at different pH values, at 100 μM substrate concentration. Background uptake rates (empty vector) were subtracted.
(c) AtPUP2-mediated adenine uptake at different substrate concentrations. Experiments were performed at pH 3.5. Background uptake rates (empty vector) were subtracted.
(d) Substrate specificity of AtPUP2. Uptake of adenine by AtPUP2 in pDR195 was determined at 60 μM adenine and pH 3.5 in presence of 10-fold excess (600 μM) of unlabeled competitors. Values represent the mean ± SD of three independent experiments; BAP, benzylaminopurine; iPA, isopentenyladenine; ZR, *trans*-zeatin riboside.

uptake was strongly inhibited by isopentenyladenine, kinetin, benzylaminopurine, *trans*- and *cis*-zeatin, and adenine. Significant inhibition was also observed with *trans*-zeatin riboside.

In contrast, AtPUP3 did not show any transport activity in yeast. The MG887-1 yeast strain expressing AtPUP3 did not take up adenine and was not able to grow on adenine as the sole N-source. In addition, AtPUP2 and 3 did not complement the growth deficiency of the adenine auxotrophic mutant DM734-238D on adenosine as the sole purine source. Differences in activity have been observed in cases of other transporters and can often be attributed to inefficient targeting of the transporters to the plasma membrane of yeast. However, the lack of complementation does not exclude that PUP3 has a different substrate or function.

### *Expression pattern of three PUPs in* Arabidopsis

Expression of the three PUP genes was examined histochemically using promoter–GUS fusions and RT-PCR with gene-specific primers. For histochemical analysis, two promoter fragments of *AtPUP1* (0.8 and 1.9 kb), two promoter fragments of *AtPUP2* (0.9 and 1.9 kb), and a 0.7 kb promoter fragment (region upstream of the translation start up to the next upstream ORF) of *AtPUP3* were fused to the GUS reporter gene in the binary vector pCB308 and introduced into *Arabidopsis* plants. RT-PCR analysis was carried out on total RNA isolated from different tissues of mature *Arabidopsis* plants grown in soil.

Thirty-seven out of 40 AtPUP1-GUS $T_1$ plants analyzed showed qualitatively the same expression pattern. Ten

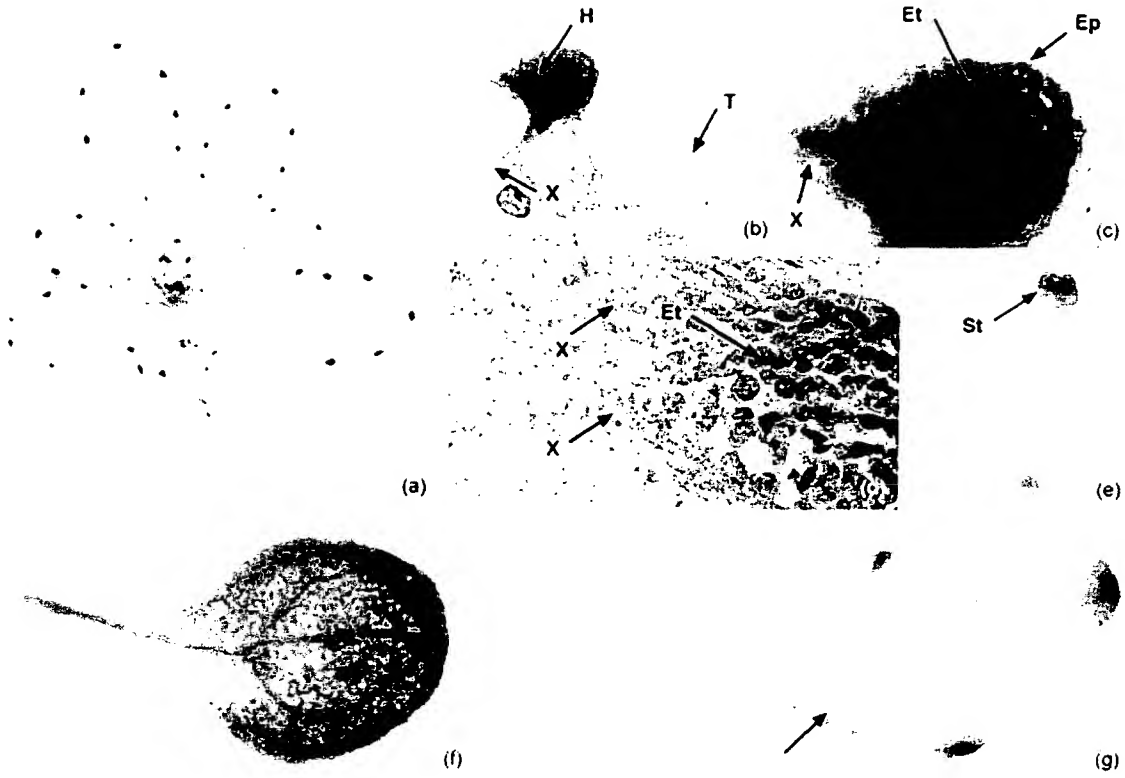



**Figure 7.** Analysis of AtPUP1 expression by promoter–GUS fusions. *Arabidopsis* plants transformed with promoter–GUS fusion constructs were stained with 1 mM X-Gluc for 24 h, de-stained and documented. Transgenic *Arabidopsis* plants harboring the 1.9 kb AtPUP1 promoter fragment driving the β-glucuronidase gene.
(a) Whole plant, (b, c) close-up of hydathode with epithem and epidermis, (d) section of hydathode, (e) close-up of a developing silique, (f) cotyledon, (g) mature leaf showing patchy staining (arrows). Ep, epidermis; Et, epithem; H, hydathode; St, stigma surface; T, trichome; X, xylem.

independent transformants per construct were analyzed histochemically for the localization of GUS activity in more detail. Although significantly stronger expression was observed in transgenic plants containing the longer promoter region, the overall expression pattern was identical for both constructs. In 3-week-old plants, GUS activity was detectable at the leaf margins in hydathodes (Figure 7a). Higher magnification of hydathodes from transgenic plants revealed GUS activity in epithem cells (Figure 7b–d) located between the xylem endings and the epidermis, cells proposed to be directly involved in the retrieval of solutes from the xylem sap (Wilson *et al.*, 1991). No staining was detected in the epidermis surrounding hydathodes. In mature plants, expression of *AtPUP1* was found at the stigma surface of siliques, potentially fulfilling a function similar to that in hydathodes (Figure 7e). In 2-week-old seedlings grown on MS medium, GUS activity was also observed in cotyledons and sometimes in patches of mature leaves (Figure 7f,g). Consistent with histochemical analysis, the RT-PCR experiments showed high-expression levels of *AtPUP1* in hydathode-enriched fractions and intercostal fields. No signal was obtained in roots as shown by both reporter gene expression RT-PCR (Figure 8). In contrast to the GUS analysis, RT-PCR analysis, however, provided evidence that the *AtPUP1* gene is also expressed in other tissues such as petioles, stems, major veins, and flowers (Figure 8). These discrepancies may be caused by differences in the status of the plants used for analysis (see patches in Figure 7g), the absence of additional promoter elements or differences in mRNA stability.

In AtPUP2-GUS plants, reporter gene activity was found in more than 14 independent $T_1$ lines and the expression pattern was indistinguishable for both constructs containing promoters differing in length. In 3- and 5-week-old plants, GUS activity was detectable in the vascular system of source leaves (Figure 9a(i)). In older plants, strong staining was also present in the vascular tissue of cauline leaves (Figure 9a(ii)). No GUS activity was observed in stems and roots. Longitudinal- and cross-sections revealed that GUS staining is limited to the phloem (Figure 9a(iii–v)), consistent with a potential role of AtPUP2 in long-distance transport of adenine and cytokinins. The results obtained by RT-PCR are consistent with the expression pattern of the *AtPUP2* promoters in the vascular tissue because AtPUP2

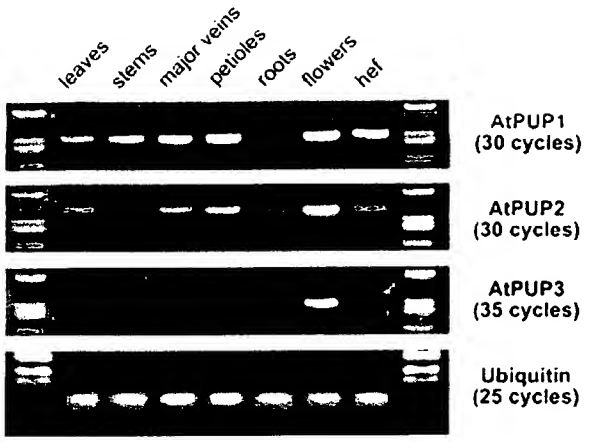



**Figure 8.** Analysis of AtPUP1–3 expression by RT-PCR. RNA from different tissues of mature *Arabidopsis* plants were converted to cDNA by reverse transcription. A 524 bp fragment of *AtPUP1* and a 619 bp fragment of *AtPUP2* were amplified by 30 PCR cycles. To obtain the 515 bp *AtPUP3* fragment, 35 PCR cycles were carried out. A 78 bp ubiquitin fragment was amplified simultaneously by 25 PCR cycles as control. hef, hydathode-enriched fraction.

transcript was detectable in major vein fractions and petioles (Figure 8). The weak signal obtained in the hydathode-enriched fraction is probably derived from contamination with other leaf tissue (veins). Both methods did not detect expression in stems. In addition, RT-PCR showed *AtPUP2* expression in flowers and roots, a pattern not observed in case of the GUS analysis, indicating potential differences in the status of the plants, the absence of additional promoter elements or differences in mRNA stability.

For AtPUP3-GUS plants, the activity of the reporter gene was detected in 12 out of 13 independent $T_1$ lines. GUS staining was restricted to pollen (Figure 9b(i–iii)), indicating a potential role of AtPUP3 in transport of purine derivatives during pollen germination and tube elongation as found in *Petunia* pollen (Kamboj and Jackson, 1987). The same expression pattern has been obtained by RT-PCR (Figure 8), confirming that the short promoter fragment contains all necessary elements.

## Discussion

### *A common transport system for adenine and cytokinins*

The presence of purines in phloem sap and cytokinins in phloem and xylem exudate strongly suggests the existence of carriers controlling long-distance transport of these substances (Beck and Wagner, 1994; Kluge and Ziegler, 1964; Weiler and Ziegler, 1981; Yong *et al.*, 2000; Ziegler and Kluge, 1962). In our report, transport studies in cell cultures indicate that adenine and cytokinins use the same transport systems. At least one set of transporters is required to be capable of cellular import *in planta*. A second set of transporters must be postulated for cellular export (Figure 10). For example, cytokinins produced in the root have to enter the xylem vessels via cellular exporters, probably from xylem parenchyma cells, with functions comparable to that of the outwardly rectifying potassium channels (Gaymard *et al.*, 1998). At the other end of the xylem, i.e. in leaves, cellular importers are required both for the uptake of cytokinins into leaf cells and for the cellular uptake and recycling of the cytokinin degradation product adenine. Furthermore, phloem loading with cytokinins, e.g. in leaves, similar to sucrose loading, requires cellular export systems and subsequently uptake carriers for the phloem cells.

The competition of adenine uptake by cytokinins in both cell cultures and yeast expressing *Arabidopsis* PUP purine permeases, demonstrated here, may indicate that cytokinins are substrates for transport. However, competition is not direct proof of the actual transport of a compound. Thus, direct uptake studies using radiolabeled cytokinins were carried out with yeast expressing AtPUP1. AtPUP1 mediates the transport of both cytokinins and adenine when expressed in yeast. AtPUP1 is assumed to function as a cellular uptake system at the plasma membrane by means of proton co-transport, because uptake is energy dependent, occurs against a concentration gradient, is stimulated by acidification, and is inhibited by protonophores (Gillissen *et al.*, 2000). When determined in yeast, the affinity of AtPUP1 for adenine was similar to the $K_m$ for *trans*-zeatin in the low micromolar range. At first sight, the affinity range may seem relatively high, considering that *trans*-zeatin serves as a phytohormone. However, the apparent affinities of *O*-glycosylating enzymes for *trans*-zeatin range from 2 to 28 μM (Dixon *et al.*, 1989; Turner *et al.*, 1987), the $K_m$ of cytokinin oxidase measured in plant extracts was found between 0.1 and 31 μM, and that of the cloned enzyme was 14 μM (Bilyeu *et al.*, 2001; Galuszka *et al.*, 1999). Thus, the affinities of enzymes involved in cytokinin metabolism and AtPUP1 are in a comparable range, further supporting the hypothesis that also *in planta*, PUPs can function in cytokinin transport. As AtPUP1 is also able to mediate adenosine transport, it may contribute to cellular uptake of adenosine, although with very low efficiency. However, in analogy to the PUP's broad selectivity, high-affinity nucleoside transporters such as ENT1,At may be candidates for such a function (Möhlmann *et al.*, 2001). Although the activity of AtPUP2 was relatively low, the overall properties regarding affinity and substrate specificity were similar to those of AtPUP1. Various cytokinin analogs including *cis*- and *trans*-zeatin, isopentenyladenine, and kinetin strongly competed for adenine uptake mediated by AtPUP1. In contrast, no activity was detected for the closely related AtPUP3, potentially resulting in ineffective plasma membrane targeting in the heterologous

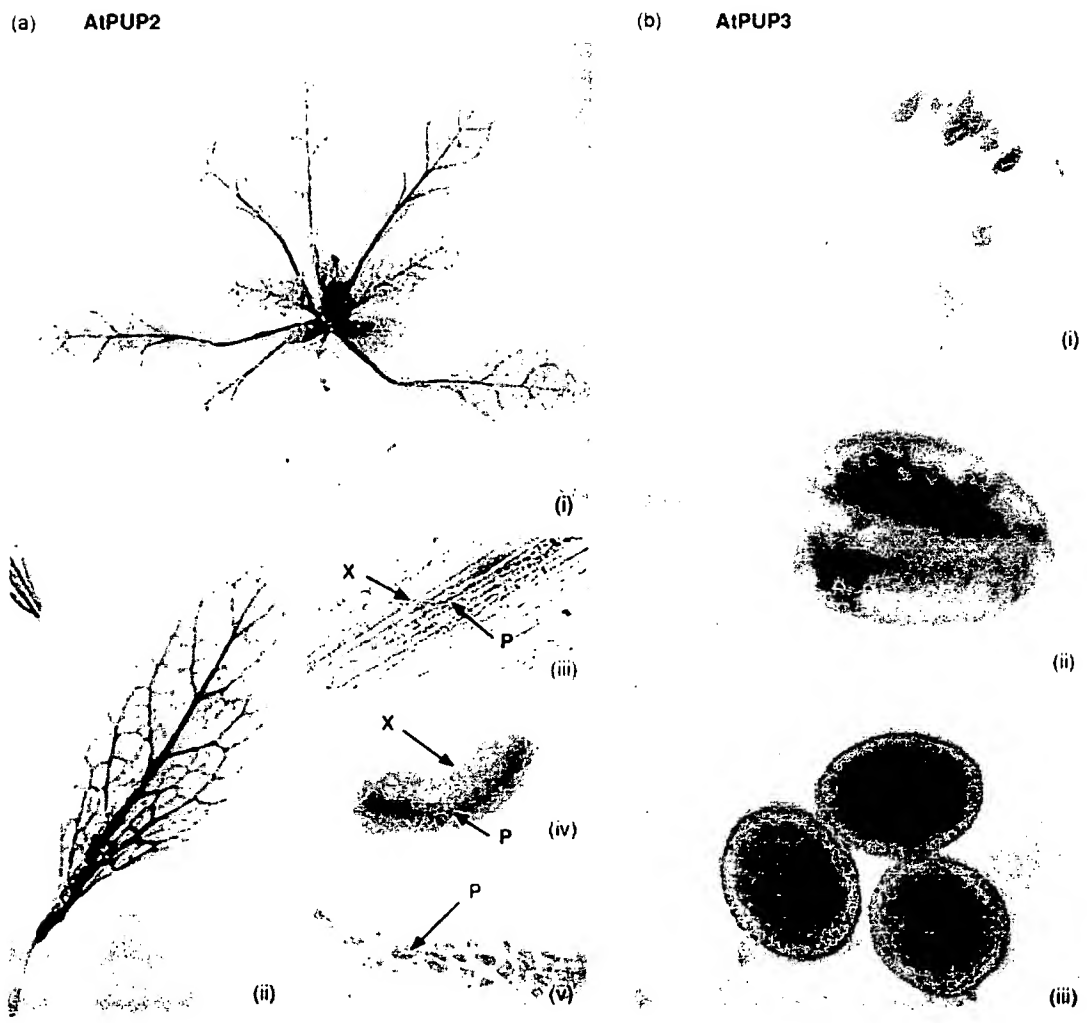



**Figure 9.** Analysis of AtPUP2 and 3 expression by promoter–GUS fusions. *Arabidopsis* plants transformed with promoter–GUS fusion constructs were stained with 1 mM X-Gluc for 24 h, de-stained and documented.
(a) Transgenic *Arabidopsis* plants harboring the 1.9 kb AtPUP2 promoter fragment driving the β-glucuronidase gene. (i) Whole 3-week-old plant showing staining of the vascular system in source leaves. (ii) Staining of the vascular tissue in a cauline leaf from a mature plant. (iii) Longitudinal section through a source leaf petiole. (iv) Cross-section of the vascular system of a source leaf. (v) Larger magnification. P, phloem; X, xylem.
(b) Transgenic *Arabidopsis* plants harboring the 0.7 kb AtPUP3 promoter fragment driving the β-glucuronidase gene. (i) Flower showing staining in anthers, (ii) anther, (iii) magnification of pollen.

expression system. Thus, in principle, it is possible that the three proteins fulfil similar functions in the plant.

### *Different roles of the PUP transporters*

As a first hint for the actual physiological function, a promoter analysis was used to determine the cellular expression pattern of the three transporters. Expression of the PUP1 promoter was confined to the epithem cells of hydathodes and to the stigma surface of siliques. A similar expression pattern had been observed for the inward $K^+$ channel AKT1 (Lagarde *et al.*, 1996). Hydathodes mediate guttation and may be involved in the retrieval solutes from the transpiration stream. With respect to phylogenesis, hydathodes are related to nectaries (Vogel, 1998). A typical hydathode consists of a particular parenchymatous tissue, the epithem, and an epidermis with modified stomata serving as water pores. The morphological features of epithem cells are large numbers of mitochondria and a lobed surface resulting in an increase of surface area facing the extracellular space, suggesting involvement not only in the active secretion of solutes, but also in the selective absorption and retrieval of both inorganic and organic solutes (Galatis, 1988; Höhn, 1951). In agreement with the findings of Wilson *et al.* (1991), the epithem is able to take up $^{14}C$-labeled aspartate fed to the leaves via the transpiration system. Similarly, nectaries, although thought to serve mainly as secretory organs, are also able to take up

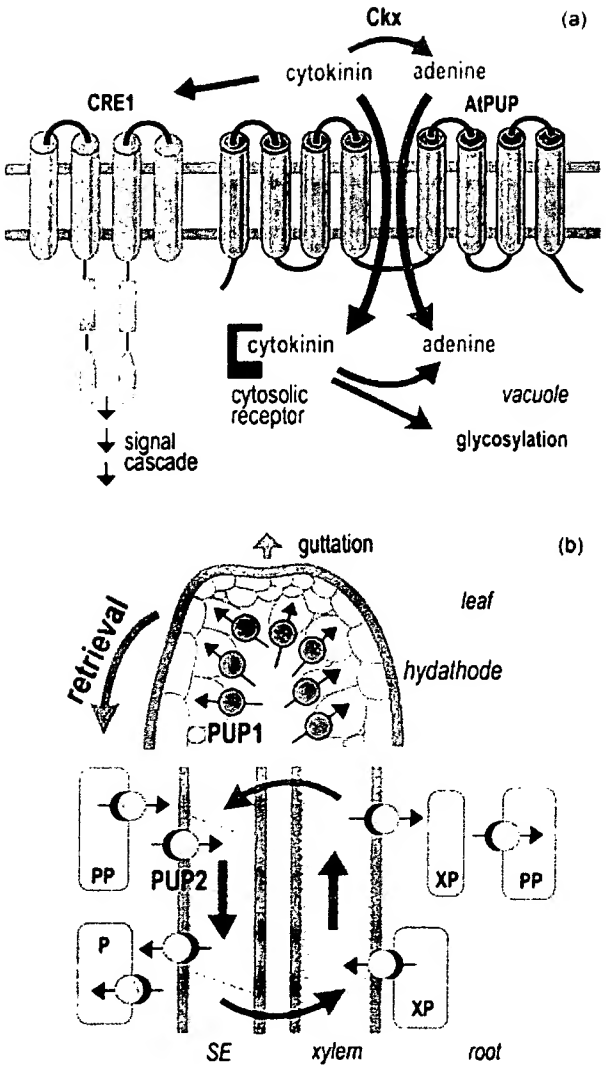



**Figure 10.** Hypothetical model of cytokinin signal perception and transport. (a) Cellular model for cytokinin metabolism and transport. The extracellular level of cytokinins is controlled by the activity of cytokinin oxidase (Ckx) and cytokinin transporter (PUP). Cytokinins can either bind to the plasma membrane receptor (CRE1) eliciting a hormone-specific signaling cascade or can be taken into the cell by a transporter, e.g. PUP1. Adenine derived from cleavage of cytokinin in the apoplasm can be retrieved into the cell by PUP1 as well. Alternatively, cytokinins that have been taken up by PUP1 may bind to an intracellular receptor, controling the same or a different signaling cascade. (b) Whole plant model for cytokinin transport. Both cellular importers and exporters are required in different cell types of the vascular system and the parenchymatous tissues surrounding them for long-distance transport of cytokinins in xylem and phloem. AtPUP1 localized in the epithem may serve as a retrieval system to preserve purines and purine derivatives including cytokinins from the guttation sap. AtPUP2 localized in the phloem of leaves, functions as a cellular import system, and thus may be involved in loading of the phloem with purines and cytokinins. SE, sieve elements; P, parenchyma, PP, phloem parenchyma; X, xylem; XP, xylem parenchyma.

radiolabeled amino acids (Ziegler and Lüttge, 1959). Therefore, the expression of AtPUP1 in the epithem of hydathodes may suggest a role in retrieval of purines and cytokinins from xylem sap. Epithem cells accumulate sulforhodamine G, indicating the presence of highly active $H^+$-ATPases and respective $H^+$-symporters for retrieval of ions like potassium and other important organic compounds (Wilson *et al.*, 1991). Consistent with a potential role in the retrieval, circumstantial evidence suggests that AtPUP1 functions as a $H^+$-co-transporter, fulfilling the requirements of a retrieval system. As AtPUP1 is able to transport another adenine analog caffeine, which is translocated in the xylem of coffee plants, a PUP1 ortholog may serve a similar function in retrieval (Mazzafera and Gonçalves, 1999). Interestingly, caffeine has also been implicated in cytokinin-like functions (Vitória and Mazzafera, 1997). PUP1 expression was also detected in cotyledons, consistent with the finding that at least in *Ricinus communis* the cotyledons take up adenine rapidly (Kombrink and Beevers, 1983). The presence of mRNA as detected by RT-PCR in other tissues and the presence of patchy GUS staining in other tissues may indicate that the gene can be induced under certain conditions or that potentially additional elements in the gene contribute to regulation.

AtPUP2 promoter expression was found in the vascular system of leaves, i.e. in the phloem of leaves, pointing to a potential role in phloem loading of adenine and cytokinins. Also in this case, RT-PCR analysis indicates that the gene may be expressed in additional tissues. The expression pattern of AtPUP3 was limited to pollen, in which it might play a role in the transport of purine derivatives during pollen germination and tube elongation (Kamboj and Jackson, 1987). It is obvious from the above discussion that multiple transporters are required for long-distance transport (Figure 10b). Thus, it is not surprising that the PUPs constitute a large, however extremely diverse, gene family, in which the most distant members share less than 20% similarity at the amino acid level. Further studies are required to obtain an overview of the transport properties, the expression pattern, and regulation of various other members of the PUP family.

Interestingly, a member of the PUP family from tomato was found as an expressed sequence tag in *Agrobacterium*-induced tumors (GenBank accession number BG131703). Similar to the situation in animals, plant tumors are also highly vascularized (Aloni *et al.*, 1995). Agrobacteria transform plant cells to tumor cells by overproducing auxins and cytokinins. The balance of cytokinins and auxins is not only critical for callus or tumor formation but also for the development of the vascular system. Despite local cytokinin overproduction, infected plants do not show systemic symptoms of cytokinin action. Further studies are required to determine whether PUPs expressed in tumors have a function in channeling cytokinins, in retrieving them or in protecting the plant from systemic effects of the infection. Similar mechanisms may also operate in the restriction of

cytokinin movement in fungal infection of leaves and may explain why local overproduction of cytokinins does not lead to systemic loss of apical dominance (Faiss *et al.*, 1997).

In summary, transport studies in *Arabidopsis* cell cultures indicate that adenine and cytokinins are transported by a common system, which has properties similar to heterologously expressed PUP transporters. Transport studies in yeast in combination with promoter reporter gene studies indicate that different members of the PUP family are involved in the long-distance transport of adenine, cytokinins, and alkaloids. Whereas AtPUP2 may play a role in phloem transport, AtPUP1 may be involved in retrieval of nucleobases and derivatives to prevent secretion in hydathodes and at the stigma surface of siliques and AtPUP3 in supply of pollen. These studies provide a basis for a more direct analysis of the physiological function of PUPs by studying the effect of specific inhibition of transport by PTGS by overexpression or by using insertional mutants.

## Experimental procedures

### *Yeast strains and growth conditions*

Yeast strains used were MG887-1 (Mat a, *fcy2, ura3*) (Gillissen *et al.*, 2000) and DM734–284D (Mat a, *ade8–18, ade2–1, arg4–16, leu2–27, trp1–1, lys2, ura3, gal7*) (DM Yeast Genetic Stock Center, Berkeley, USA). Yeast cells were cultured in YPD or minimal media. The minimal medium was supplemented with 7.4 mM adenine as the sole nitrogen source for growth analysis of MG887-1. For growth experiments with DM734–284D, 5 g $l^{-1}$ ammonium sulfate and 150 μM of the purine source were added to the minimal medium. To study caffeine toxicity in MG887-1, the minimal medium was supplemented with caffeine (0.1 and 0.2%) and 5 g $l^{-1}$ ammonium sulfate.

### *DNA work*

To generate pDR195/AtPUP1, *Not*I-restricted full-length cDNA encoding AtPUP1 was isolated from the pFL61/AtPUP1 (Gillissen *et al.*, 2000) and ligated into pDR195 (Rentsch *et al.*, 1995). Orientation of inserts was tested by restriction analysis.

The cDNA of *AtPUP2* was obtained by RT-PCR (RETROscript Kit, Ambion, Austin, USA) with RNA from *Arabidopsis thaliana* roots (ecotype Col-0) as a template. For the first-strand synthesis, random decamer primers were used. The following PCR was performed with gene-specific AtPUP2 primers (5′-GAGAGGATCCGAAGATGAAGATGAAGACAG-3′ and 5′-GAGAGGATCCGCTTTTAAGCTA-CATAATCAG-3′). The amplified *AtPUP2* cDNA was ligated with pCRII-TOPO (Invitrogen, Groningen, Netherlands) and subcloned into the yeast expression vector pDR195, previously cut with *Bam*HI and *Not*I (Rentsch *et al.*, 1995). The cDNA encoding *AtPUP3* was amplified by RT-PCR from RNA of *A. thaliana* (ecotype Col-0) flowers. Oligo dT primers were used for the first-strand synthesis. The following PCR was carried out using primers: 5′-TATCTTGGATCCAGACAAGAATGGTGAAGGCTCTTG-3′ and 5′-ATATGCGGATCCGAGTTTTAACACTCACTTAATGGGTC-3′. The obtained PCR product was digested with *Bam*HI and ligated into *Bam*HI-digested pDR195 (Rentsch *et al.*, 1995). The cDNAs encoding *AtPUP2* and *3* were sequenced, and no differences to the genomic sequence were found except the presence of an intron. The pDR195 plasmids containing the AtPUP1-, AtPUP2- or AtPUP3-coding sequence were introduced into MG887-1 and DM734–284D yeast strains following a modified method previously described (Dohmen *et al.*, 1991). Colonies able to grow on plates lacking uracil were used for further growth analysis and uptake experiments.

The promoter regions of *AtPUP1–3* were isolated by PCR with genomic DNA from *A. thaliana* ecotype Col-0 as a template. The 0.8 kb promoter fragment of *AtPUP1* was obtained using Pfu polymerase (Stratagene, La Jolla, USA), the F1a (5′-GAGATCTAGACCTAGCGGGGTAGAAACCTTG-3′) and R1 (5′-GAGAGGATCCTTCTTCTTGCTGCTTGCTG-3′) primers. The 1.9 kb promoter fragment of *AtPUP1* was isolated using Taq polymerase (Roche Molecular Biochemicals, Germany), F1b (5′-GAGATCTAGAGTCATCAAAGATTTCCTAAAC-3′) and R primers. The obtained DNA promoter fragments were subcloned into pCRII-TOPO, afterwards excised with *Bam*HI and *Xba*I and fused to the GUS gene located on the binary vector pCB308 (Xiang *et al.*, 1999). The 0.9 kb promoter fragment of *AtPUP2* was amplified using the Pfu polymerase, the F2a (5′-GAGAACTAGTGGAAACTCATAACTCCGATG-3′) and R2 (5′-GAGAGGATCCTTCTTCTTCTTGCTATAACC-3′) primers. The 1.9 kb fragment of *AtPUP2* was isolated using the Taq polymerase, the F2a primer (5′-GAGAACTAGTGATGGAACGTCTACGAACAC-3′) and R2 primers. The 0.7 kb promoter fragment of *AtPUP3* was amplified using the Pfu polymerase, the F3 (5′-GAGAACTAGTGTTCTTGTGTATAAGTAATG-3′) and R3 (5′-GAGAGGATCCTTGTCTGATTTCTTGTGG-3′) primers. The PCR products were cut with *Spe*I and *Bam*HI and ligated into pCB308. Sequence analysis of the inserts revealed five point mutations in the longer promoter fragment of *AtPUP1* that did not affect the qualitative expression pattern. Transformation of *A. tumefaciens* PGV2260 using the resulting pCB308 derivatives was performed as described by Deblaire *et al.* (1985).

### *Plant transformation and histochemical analysis of GUS activity*

*Arabidopsis thaliana* L. Heynh. ecotype Col-0 plants were transformed via vacuum infiltration (Clough and Bent, 1998). Transgenic *Arabidopsis* plants were selected with BASTA herbicide (Aventis Crop Science, Frankfurt, Germany). Histochemical assays for β-glucuronidase activity in transgenic plants were performed as described in Martin *et al.* (1992). Tissues were cut into 2 mm × 5 mm pieces, incubated in GUS staining solution containing 100 mM sodium phosphate (pH 7.0), 10 mM EDTA, 3 mM $K_4[Fe(CN)_6]$, 0.5 mM $K_3[Fe(CN)_6]$, 0.1% (v/v) Triton X-100, and 2 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-gluc) for 3–24 h at 37°C. Slight vacuum was applied before incubation to facilitate substrate infiltration. For resin sections, X-gluc stained tissues were fixed in 4% glutaraldehyde, 50 mM sodium phosphate (pH 7.0) overnight at 4°C. Fixed tissues were dehydrated in EtOH and embedded in LR White resin (London Resin Company Ltd, Berkshire, UK). Embedded material was cut into 1.5–5 μm sections with glass knives using an ultramicrotome and observed by bright field, phase contrast, and dark field microscopy. For the fresh sections, tissues were embedded in 5% low melting agarose in 50 mM sodium phosphate buffer pH 7.0. After solidification, agar blocks were cut, and fresh sections (75–150 μm) were made with razor blades using a vibratome. Sections were collected in ice-cold water and incubated in GUS staining solution described above, but without 0.5 mM $K_3[Fe(CN)_6]$. For chlorophyll-containing tissues, fresh sections were cleared in 70% EtOH.

### *RT-PCR analysis*

RNA for RT-PCR analysis was extracted from suspension cells and mature soil-grown *Arabidodpsis* plants. Major veins, the regions around leaf edges containing hydathodes (hydathode-enriched fraction), and intercostal fields were excised from leaves, and immediately frozen in liquid nitrogen. Aliquots of 2 μg RNA were used as a template for the first strand synthesis, using RETROscript Kit (Ambion, Austin, USA) according to the manufacturer's instructions. An aliquot of 2 μl of the first strand cDNA was used for PCR with gene-specific primers. To avoid amplification of genomic DNA, reverse primers were spanned splice sites. The 524 bp AtPUP1 fragment was amplified by 30 PCR cycles using primers: 5′-CTAACAACGCGGAAAACAAGC-3′ and 5′-CTCTTGCTATCACCTTAAAATCTC-3′. The 619 bp AtPUP2 transcript was obtained by 30 PCR cycles with specific primers: 5′-TATCTTGGTACCAAAGGATCTGGTTTCCAAGC-3′ and 5′-TCCTGCTATCACCTTGAAATCG-3′. The 515 bp AtPUP3 fragment was amplified by 35 PCR cycles with primers: 5′-ACAATGTGGGTGATAGTACAAG-3′ and 5′-CTTTGGTAAGGCCTTGAAAATC-3′. To ensure that equal amounts of cDNA were added to each PCR reaction, a cDNA fragment of the constitutively expressed ubiquitin gene was amplified simultaneously by 25 PCR cycles using primers: 5′-GAATCCACCCTCCACTTGGTC-3′ and 5′-CGTCTTTCCCGTTAGGGTTTT.

### *Transport measurements into* Arabidopsis *suspension cells*

The suspension cell culture from *Arabidopsis thaliana* ecotype Landsberg erecta (May and Leaver, 1993) was a gift from Mike Bevan (The Jones Innes Centre, Norwich, UK). Cells were maintained as described by Fuerst *et al.* (1996). For the uptake experiments, cells were harvested 4 days after dilution into new medium by centrifugation at 18 *g* for 3 min, and were re-suspended in fresh cytokinin-free medium containing 10 mM potassium phosphate buffer (pH 5.7) at a concentration of 0.25 ml packed cells per 1 ml of suspension. To start the reaction, 850 μl of this suspension was mixed with 50 μl medium containing $^{14}$C-adenine (final concentration 13 Bq μl$^{-1}$) and the unlabeled analog (final concentration 20 μM). Samples of 170 μl were removed after 1–4 min, filtered on glass-fiber filters, and washed twice with 10 ml medium. Radioactivity on the filters was determined by liquid scintillation spectrometry (Wallac, Turku, Finland). The protonophore CCCP was added together with the radiolabeled mix. Competition experiments were performed with a 10-fold excess of the respective competitor resulting in a final concentration of 200 μM.

### *Transport measurements into yeast cells*

For adenine uptake studies, yeast cells were harvested at $OD_{600}$ of 0.6, washed in water and re-suspended in 100 mM sodium citrate buffer (pH 3.5) containing 1% glucose to a final $OD_{600}$ of 12. An aliquot of 100 μl of the cell suspension was pre-incubated for 2 min at 30°C. To start the reaction, a 100 μl buffer containing 13 Bq μl$^{-1}$ $^{14}$C-labeled adenine (Amersham), 1% glucose, and the unlabeled analog, as indicated, was added. Samples of 50 μl were removed after 30, 60, 120 and 180 sec, transferred to 4 ml ice-cold water, filtered on glass-fiber filters, and then washed with 8 ml water before radioactivity was determined. For analysis of the pH dependence, cells were washed in water, and then re-suspended in 100 mM sodium phosphate buffer at different pH. For *trans*-zeatin uptake assays, yeast cells were harvested at $OD_{600}$ of 0.6, washed, and re-suspended in 100 mM sodium phosphate buffer (pH 5.0) to a final $OD_{600}$ of 10. An aliquot of 500 μl of the cell suspension was pre-incubated for 3 min at 30°C, supplemented with glucose (final concentration 10 mM) and incubated for a further 2 min. To start the reaction, a 39.5 μl buffer containing labeled *trans*-2-$^{3}$H-zeatin (final concentration 87.3 Bq μl$^{-1}$; Olchemlm, Olomouc, Czech Republic) and unlabeled analogs was added (final concentration 100 μM). Samples of 100 μl were removed after 20, 40, 60, and 90 sec, transferred to 4 ml ice-cold 5 mM adenine solution, filtered on glass-fiber filters, and washed with 8 ml adenine (5 mM). Transport measurements were repeated independently and represent the mean of at least three experiments. For analysis of purity, 100 μl of a radioactive sample (1 : 200 dilution) was loaded (using an isocratic gradient: methanol (0.2 ml min$^{-1}$) and 50 mM acetic acid/ammonia (0.6 ml min$^{-1}$)) on a HPLC column (LUNA 5 μm C18(2), Phenomenex, Germany). The chromatograms were recorded by UV detection (diode array, Kontron, Germany) or radioactivity (LB507B, Berthold, Germany). Retention times of *trans*-zeatin and adenine were 13.91 and 4.63 min, respectively.

## Acknowledgements

We are grateful to Bettina Stadelhofer (Analytics Facility ZMBP) and Bettina Million for their excellent technical assistance and we would like to thank Thomas Schmülling (Free University Berlin) and Felicity de Courcy for their critical reading of the manuscript. We gratefully acknowledge the support of the Deutsche Forschungsgemeinschaft (SPP '$CO_2$ and transport' and SFB 446).

## Supplementary Material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/TPJ/TPJ1700/TPJ1700sm.htm

**Figure S1.** Structure and radiochemical purity of *trans*-$^{3}$H-zeatin.
(a) Structural formula of *trans*-$^{3}$H-zeatin.
(b) HPLC profile of *trans*-$^{3}$H-zeatin using a radioactivity detector. The retention times of adenine (4.63 min) and *trans*-zeatin (13.91 min) are indicated.

**Figure S2.** Phylogenetic tree of *Arabidopsis* PUP family. Maximum parsimony analysis was performed using PAUP version 4.0b8a (Swofford, 1998) with all DNA characters unweighted and gaps scored as missing characters. Heuristic tree searches were executed using 100 random sequence additions and the tree bisection re-connection branch swapping algorithm with random sequence analysis. The complete alignment was based on 364 sites. A total of 313 sites were phylogenetically informative. The AtPUP paralogs were re-named and grouped into four clades. The alignment underlying the phylogenetic tree is available at http://www.uni-tuebingen.de/plantphys/~PUPtree.

**Figure S3.** Alignment of AtPUP1–3. The deduced amino acid sequences were aligned by using the MEGALIGN program (DNASTAR, Madison, WI). Identical amino acids are shaded.

**Figure S4.** Prediction of 10 putative membrane spanning regions in the amino acid sequence of AtPUP1–3. Hydropathy plots were performed by using THMM1.0 (Sonnhammer *et al.*, 1998). N- and C-termini were predicted to be cytosolic.

## References

**Aloni, R., Pradel, K.S. and Ullrich, C.I.** (1995) The three-dimensional structure of vascular tissues in *Agrobacterium tumefaciens*-induced crown galls and in the host stems of *Ricinus communis* L. *Planta*, **196**, 597–605.

**Astot, C., Dolezal, K., Nordstrom, A., Wang, Q., Kunkel, T., Moritz, T., Chua, N.H. and Sandberg, G.** (2000) An alternative cytokinin biosynthesis pathway. *Proc. Natl. Acad. Sci. USA*, **97**, 14778–14783.

**Bard, M., Neuhauser, J.F. and Leeds, N.D.** (1980) Caffeine resistance of *Saccharomyces cerevisiae*. *J. Bacteriol.* **141**, 999–1002.

**Beck, E. and Wagner, B.M.** (1994) Quantification of the daily cytokinin transport from the root to the shoot of *Urtica dioica* L. *Bot. Acta*, **107**, 342–348.

**Beveridge, C.A., Murfert, I.C., Kerhoas, L., Sotta, B., Miginiac, E. and Rameau, C.** (1997) The shoot controls zeatin riboside export from pea roots. Evidence from the branching mutant *rms4*. *Plant J.* **11**, 339–345.

**Bilyeu, K.D., Cole, J.L., Laskey, J.G., Riekhof, W.R., Esparza, T.J., Kramer, M.D. and Morris, R.O.** (2001) Molecular and biochemical characterization of a cytokinin oxidase from maize. *Plant Physiol.* **125**, 378–386.

**Brault, M., Maldiney, R. and Miginiac, E.** (1997) Cytokinin-binding proteins. *Physiol. Plant.* **100**, 520–527.

**Clough, S.J. and Bent, A.F.** (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

**Deblaire, R., Bytebier, B., Greve, H., Deboeck, F., Schell, J., van Montgau, M. and Leemans, J.** (1985) Efficient octopine Ti plasmid derived vectors for *Agrobacterium*-mediated gene transfer to plants. *Nucl. Acids Res.* **13**, 4777–4788.

**Desimone, M., Catoni, E., Ludewig, U., Hilpert, M., Schneider, A., Kunze, R., Tegeder, M., Frommer, W.B. and Schumacher, K.** (2002) A novel superfamily of transporters for allantoin and other oxo derivatives of nitrogen heterocyclic compounds in *Arabidopsis*. *Plant Cell*, **14**, 847–856.

**Detke, S.** (1998) Cloning of the *Candida albicans* nucleoside transporter by complementation of nucleoside transport-deficient *Saccharomyces*. *Yeast*, **14**, 1257–1265.

**Dixon, S.C., Martin, R.C., Mok, M.C., Shaw, G. and Mok, D.W.S.** (1989) Isolation of *O*-glucosyltransferase from *P. lunatus* and comparison to *O*-xylosyltransferase from *P. vulgaris*. *Plant Physiol.* **90**, 1316–1321.

**Dohmen, R.J., Straffer, A.W.M., Honer, C.B. and Hollenberg, C.P.** (1991) An efficient transformation procedure enabling long-term storage of competent cells of various yeast genera. *Yeast*, **7**, 691–692.

**Emery, R.J.N., Ma, Q. and Atkins, C.A.** (2000) The forms and sources of cytokinins in developing white lupine seeds and fruits. *Plant Physiol.* **123**, 1593–1604.

**Faiss, M., Zalubilova, J., Strnad, M. and Schmulling, T.** (1997) Conditional transgenic expression of the *ipt* gene indicates a function for cytokinins in paracrine signaling in whole tobacco plants. *Plant J.* **12**, 401–415.

**Fuerst, R.A.U.A., Soni, R., Murray, A.H. and Lindsey, K.** (1996) Modulation of cyclin transcript levels in cultured cells of *Arabidopsis thaliana*. *Plant Physiol.* **112**, 1023–1033.

**Fußeder, A., Ziegler, P., Peters, W. and Beck, E.** (1989) Turnover of *O*-glucosides of dihydrozeatin and dihydrozeatin-9-riboside during the cell growth cycle of photoautotrophic cell suspension cultures of *Chenopodium rubrum*. *Bot. Acta*, **102**, 335–340.

**Galatis, B.** (1988) Microtubules and epithem-cell morphogenesis in hydathodes of *Pilea cadierei*. *Planta*, **176**, 287–297.

**Galuszka, P., Frebort, I., Sebala, M., Strand, M. and Pec, P.** (1999) Cytokinin oxidase: the key enzyme in the biodegradation of cytokinins. In *Advances in Regulation of Plant Growth and Development* (Strand, M., Pec, P. and Beck, E., eds). Praque: PERES publishers, pp. 39–48.

**Gaymard, F., Pilot, G., Lacombe, B., Bouchez, D., Bruneau, D., Boucherez, J., Michaux-Ferriere, N., Thibaud, J.B. and Sentenac, H.** (1998) Identification of a plant shaker-like outward channel involved in $K^+$ release into the xylem sap. *Cell*, **94**, 647–655.

**Gillissen, B., Bürkle, L., André, B., Kühn, C., Rentsch, D., Brandl, B. and Frommer, W.B.** (2000) A new family of high-affinity transporters for adenine, cytosine and purine derivatives in *Arabidopsis*. *Plant Cell*, **12**, 291–300.

**Hannan, M.A. and Nasim, A.** (1977) Caffeine enhancement of radiation killing in different strains of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **158**, 111–116.

**Höhn, K.** (1951) Beziehungen zwischen Blutung und Guttation bei *Zea mays*. *Planta*, **39**, 65–74.

**Houba-Hérin, N., Plethe, C., d'Alayer, J. and Lalou, M.** (1999) Cytokinin oxidase from *Zea mays*: purification, cDNA cloning and expression in moss protoplasts. *Plant J.* **17**, 615–626.

**Hwang, T. and Sheen, J.** (2001) Two-component circuitry in *Arabidopsis* cytokinin signal transduction. *Nature*, **413**, 383–389.

**Inoue, T., Higuchi, M., Hashimoto, Y., Seki, M., Kobayashi, M., Kato, T., Tabata, S., Shinozaki, K. and Kakimoto, T.** (2001) Identification of CRE1 as a cytokinin receptor from *Arabidopsis*. *Nature*, **409**, 1060–1063.

**Kakimoto, T.** (1996) CKI1, a histidine kinase homolog implicated in cytokinin signal transduction. *Science*, **274**, 982–985.

**Kakimoto, T.** (2001) Identification of plant cytokinin biosynthetic enzymes as dimethylallyl diphosphate: atp/adp isopentenyltransferases. *Plant Cell Physiol.* **42**, 677–685.

**Kamboj, R.K. and Jackson, J.F.** (1987) Purine nucleoside transport in petunia pollen is an active, carrier-mediated system not sensitive to nitrobenzylthioinosine and not renewed during pollen tube growth. *Plant Physiol.* **84**, 688–691.

**Kluge, M. and Ziegler, H.** (1964) Der ATP-Gehalt der Siebröhrensäfte von Laubbäumen. *Planta*, **61**, 167–177.

**Kombrink, E. and Beevers, H.** (1983) Transport of purine and pyrimidine bases and nucleosides from endosperm to cotyledons in germinating castor bean seedlings. *Plant Physiol.* **73**, 370–376.

**Kulaeva, O.N., Karavaiko, N.N., Selivankina, S.Y., Zemlyachenko, Y.V. and Shipilova, S.V.** (1995) Receptor of *trans*-zeatin involved in transcription activation by cytokinin. *FEBS Lett.* **366**, 26–28.

**Kulaeva, O.N., Zagranichnaya, T.K., Brovko, F.A., Karavaiko, N.N., Selivankina, S.Y., Zemlyachenko, Y.V., Hall, M., Lipkin, V.M. and Boziev, K.M.** (1998) A new family of cytokinin receptors from cereals. *FEBS Lett.* **423**, 239–242.

**Lagarde, D., Basset, M., Lepetit, M., Conejero, G., Gaymard, F., Astruc, S. and Grignon, C.** (1996) Tissue-specific expression of *Arabidopsis AKT1* gene is consistent with a role in $K^+$ nutrition. *Plant J.* **9**, 195–203.

**Ludewig, F. and Sonnewald, U.** (2000) High $CO_2$-mediated down-regulation of photosynthetic gene transcripts is caused by accelerated leaf senescence rather than sugar accumulation. *FEBS Lett.* **479**, 19–24.

**Martin, R.C., Mok, M.C., Habben, J.E. and Mok, D.W.S.** (2001) A maize cytokinin gene encoding an *O*-glucosyltransferase specific to *cis*-zeatin. *Proc. Natl. Acad. Sci. USA*, **98**, 5922–5926.

**Martin, R.C., Mok, M.C. and Mok, D.W.S.** (1999a) Isolation of a cytokinin gene, *ZOG1*, encoding zeatin *O*-glucosyltransferase from *Phaseolus lunatus*. *Proc. Natl. Acad. Sci. USA*, **96**, 284–289.

**Martin, R.C., Mok, M.C. and Mok, D.W.S.** (1999b) A gene encoding the cytokinin enzyme zeatin *O*-xylosyltransferase of *Phaseolus vulgaris*. *Plant Physiol.* **120**, 553–557.

**Martin, T., Wöhner, R.V., Hummel, S., Willmitzer, L. and Frommer, W.B.** (1992) The GUS reporter system as a tool to study plant gene expression. In *GUS Protocols: Using the GUS Gene as a*

*Reporter of Gene Expression* (Gallagher, S.R., ed.). San Diego: Academic Press, pp. 23–43.

**Mäser, P., Sütterlin, C., Kralli, A. and Kaminsky, R.** (1999) A nucleoside transporter from *Trypanosoma brucei* involved in drug resistance. *Science*, **285**, 242–244.

**May, M. and Leaver, C.** (1993) Oxidative stimulation of glutathione synthesis in *Arabidopsis thaliana* suspension cultures. *Plant Physiol.* **103**, 621–627.

**Mazzafera, P. and Gonçalves, K.V.** (1999) Nitrogen compounds in the xylem sap of coffee. *Phytochemistry*, **50**, 383–386.

**Möhlmann, T., Mezher, Z., Schwerdtfeger, G. and Neuhaus, H.E.** (2001) Characterisation of a concentrative type of adenosine transporter from *Arabidopsis thaliana* (ENT1,At). *FEBS Lett.* **509**, 370–374.

**Mok, D.W. and Mok, M.C.** (2001) Cytokinin metabolism and action. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 89–118.

**Rentsch, D., Laloi, M., Rouhara, I., Schmelzer, E., Delrot, S. and Frommer, W.B.** (1995) NTR1 encodes a high affinity oligopeptide transporter in *Arabidopsis*. *FEBS Lett.* **370**, 264–268.

**Sonnhammer, E.L.L., von Heijne, G. and Krogh, A.** (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. In *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology* (Glasgow, J., Littlejohn, T., Major, F., Lathrop, R., Sankoff, D. and Sensen, C., eds). Menlo Park, CA: AAAI Press, pp. 175–218.

**Suzuki, T., Miwa, K., Ishikawa, K., Yamada, H., Aiba, H. and Mizuno, T.** (2001) The *Arabidopsis* sensor His-kinase, AHk4, can respond to cytokinins. *Plant Cell Physiol.* **42**, 107–113.

**Swarup, R., Marchant, A. and Bennett, M.J.** (2000) Auxin transport: providing a sense of direction during plant development. *Biochem. Soc. Trans.* **28**, 481–485.

**Swofford, D.L.** (1998) PAUP* 4.0 beta version. In **Phylogenetic Analysis Using Parsimony (and Other Methods)*. Sunderland, MA: Sinauer Associates.

**Takei, K., Sakakibara, H. and Sugiyama, T.** (2001) Identification of genes encoding adenylate isopentenyltransferase, a cytokinin biosynthesis enzyme. *Arabidopsis thaliana*. *J. Biol. Chem.* **276**, 26405–26410.

**Turner, J.E., Mok, D.W.S., Mok, M.C. and Shaw, G.** (1987) Isolation and purification of an enzyme catalyzing the formation of *O*-xylosylzeatin in *Phaseolus vulgaris* embryos. *Proc. Natl. Acad. Sci. USA*, **84**, 3714–3717.

**Ueguchi, C., Sato, S., Kato, T. and Tabata, S.** (2001) The AHK4 gene involved in the cytokinin-signaling pathway as a direct receptor molecule in *Arabidopsis thaliana*. *Plant Cell Physiol.* **42**, 751–755.

**Vitória, A.P. and Mazzafera, P.** (1997) Cytokinin-like effects of caffeine in bioassays. *Biol. Plant.* **40**, 329–335.

**Vogel, S.** (1998) Remarkable nectaries: structure, ecology, organophyletic perspectives IV. Miscellaneous cases. *Flora*, **193**, 225–248.

**Weiler, E. and Ziegler, H.** (1981) Determination of phytohormones in the phloem exudate from tree species by radioimmunoassay. *Planta*, **152**, 168–170.

**Werner, T., Motyka, V., Strnad, M. and Schmülling, T.** (2001) Regulation of plant growth by cytokinin. *Proc. Natl. Acad. Sci. USA*, **98**, 10487–10492.

**Wilson, T.P., Canny, M.J. and McCully, M.E.** (1991) Leaf teeth, transpiration and the retrieval of apoplastic solutes in balsam poplar. *Physiol. Plant.* **83**, 225–232.

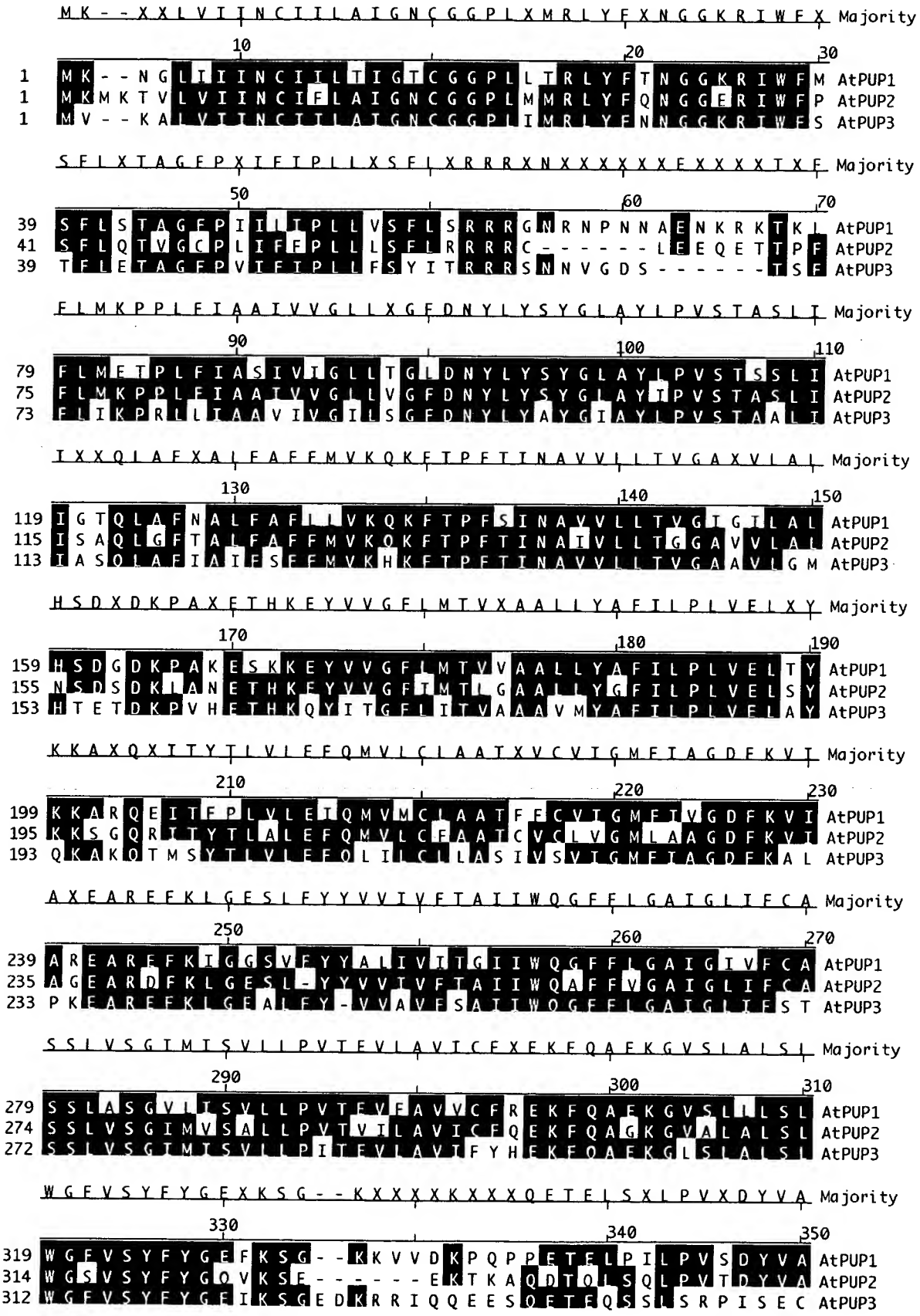
**Xiang, C., Han, P., Lutziger, I., Wang, K. and Oliver, D.J.** (1999) A mini binary vector series for plant transformation. *Plant Mol. Biol.* **40**, 711–717.

**Yong, J.W., Wong, S.C., Letham, D.S., Hocart, C.H. and Farquhar, G.D.** (2000) Effects of elevated [$CO_2$] and nitrogen nutrition on cytokinins in the xylem sap and leaves of cotton. *Plant Physiol.* **124**, 767–780.

**Ziegler, H. and Kluge, M.** (1962) Die Nucleinsäuren und ihre Bausteine im Siebröhrensaft von *Robinia Pseudoacacia* L. *Planta*, **5**, 144–153.

**Ziegler, H. and Lüttge, U.** (1959) Über die Resorption von $C^{14}$-Glutaminsäure durch sezernierende Nektarien. *Naturwissenschaften*, **5**, 176–177.

# Suppl. Fig. 3, Bürkle et al.




Decoration 'Decoration #1': Shade (with solid black) residues that match the Consensus exactly.

*Protein Science* (1993), *2*, 2217–2225. Cambridge University Press. Printed in the USA.


# Development of an in vivo method to identify mutants of phage T4 lysozyme of enhanced thermostability

PHILIP PJURA,[1] MASAZUMI MATSUMURA,[2] WALTER A. BAASE, AND BRIAN W. MATTHEWS

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, University of Oregon, Eugene, Oregon 97403

(RECEIVED July 29, 1993; ACCEPTED September 16, 1993)

**Abstract**

An M13 bacteriophage-based in vivo screening system has been developed to identify T4 lysozyme mutants of enhanced thermal stability. This system takes advantage of easy mutagenesis in an M13 host, the production of functional T4 lysozyme during M13 growth, and the ability to detect lysozyme activity on agar plates. Of several mutagenesis procedures that were tested, the most efficient was based on misincorporation by avian myeloma virus reverse transcriptase. This one-step mutagenesis and screening system has been used to find 18 random single-site mutant lysozymes, of which 11 were heat resistant. Each of these had a melting temperature within 0.8–1.4 °C of wild type, suggesting that the screening system is quite sensitive.

**Keywords:** mutagenesis; protein stability; screening

Following the pioneering studies of Streisinger et al. (1961), Alber and Wozniak (1985) attempted to develop a convenient method for generating and locating thermostable mutants of T4 lysozyme using the T4 phage system itself. The screening method was based on the ability to detect lysozyme activity as a digestion halo surrounding phage plated on a bacterial lawn, following exposure to chloroform (Streisinger et al., 1961, 1966). One of the initial mutants that was characterized was subsequently identified as Cys 54 → Tyr (Alber & Matthews, 1987). In this case the apparent heat resistance may have been due to elimination of one of the chemically reactive cysteines rather than a thermodynamic effect per se.

Because of the obvious benefits of a system that would allow rapid, in vivo screening of large numbers of mutants for stability, an attempt has been made to develop a similar, phage-based system for mutagenesis and screening of T4 lysozyme, but using instead the single-strand phage M13mp18. Because M13 does not produce lysozyme activity of its own, but will express a T4 lysozyme gene cloned into it, this activity can be visualized using the same halo assay used for T4 phage. In addition, because M13 is a small, single-strand phage, mutagenesis, screening, and sequencing of the gene can be carried out quickly, efficiently, and easily, without the need for intermediate cloning steps.

Methods to identify thermostable variants have previously been developed for kanamycin nucleotidyltransferase (Matsumura & Aiba, 1985; Liao et al., 1986), subtilisin (Bryan et al., 1986), and glucose dehydrogenase (Makino et al., 1989).

## Results

### *Plating assays*

The halo assay is based on the fact that, like T4 phage, the M13 vector containing the T4e gene expresses T4 lysozyme activity when plated on bacterial lawns and this activity can be visualized as clear circles, or "haloes," surrounding the phage plaques when the plates are treated with chloroform vapor (Streisinger et al., 1961). When the assay plates are incubated at low temperatures, all mutants and the wild-type protein exhibit haloes of approximately the same size. As the incubation temperature is increased, the haloes increase in size uniformly and the proteins remain undistinguishable, until the temperature reaches a certain critical value, which varies from mutant to mutant, above which the halo size of that mutant rap-

Reprint requests to: Brian W. Matthews, Institute of Molecular Biology, Howard Hughes Medical Institute, University of Oregon, Eugene, Oregon 97403.

[1] Present address: Department of Chemical Engineering, University of Delaware, Colburn Laboratory, Newark, Delaware 19716.

[2] Present address: Supragen, Inc., 1670 Pierce Street, Lakewood, Colorado 80214.

idly decreases. The critical temperature depends on the reversible melting temperature of the protein. By carefully adjusting the incubation temperature it is possible to distinguish mutants from a reference protein for which the melting temperature differs by less than 2 °C (Fig. 1).

When using WT* as the reference protein (see Materials and methods), plates were incubated for 6 h at 37 °C, followed by 6 h at 42 °C, before exposure to chloroform overnight for ~10–12 h. WT*-containing plaques produced sharp haloes with an average diameter 7 mm, whereas phage containing the mutant N144E/WT*, which has an increase in melting temperature of 1.5 °C at pH 6.5 (Daopin et al., 1991a), produced haloes of 8–9 mm. When using H31N/D70N as reference, a 12-h incubation at 37 °C was used. This produced haloes of 2–2.5 mm average diameter, compared to a 5-mm average diameter halo for WT*.

*Hydroxylamine mutagenesis*

Treatment of the phage with hydroxylamine for periods ranging from 1 to 8 h gave stocks with surviving titers of 15 to 0.005% of the starting value. In total, approximately 20,000 plaques were examined, of which 49 were selected and phage stocks made for rescreening. Of these, 10 were sequenced and were found to code for WT* lysozyme. On the assumption that mutations elsewhere in the phage may have caused increased levels of expression of the enzyme, one of the putative increased-expression mutations of M13 was used as the cloning vehicle. Also, the background was changed from WT* to H31N/D70N. Thirty-eight thousand plaques were screened, of which 20 were picked for rescreening. Of these, four were sequenced. All four coded for the mutant Asp 41 → Val, which in the wild-type background has an increase in melting temperature, $\Delta T_m$, of 0.6 °C (Table 1).

*Thionucleotide misincorporation mutagenesis*

Mutagenesis was carried out in the H31N/D70N background. Approximately 3,000–4,000 plaques were screened, of which 29 candidates were picked for rescreening. Of these, eight were judged to be promising after a second round of screening and were sequenced. Four of these were identified as the pseudo-revertant with lysine at position 70 (H31N/D70K), which is 1.7 °C more stable at pH 5.4 than its parent (H31N/D70N) (Table 1). The other four coded for the parent enzyme.

*Avian myeloma virus (AMV) reverse transcriptase misincorporation mutagenesis*

In a third series of experiments, a variation of the thionucleotide misincorporation method (Fig. 2) was used to try to further increase the mutagenic efficiency. This uses AMV reverse transcriptase instead of the Klenow enzyme for the misincorporation step. This enzyme lacks the proofreading activity of Klenow, and so it is unable to excise a misincorporated nucleotide once it has been inserted into the extended primer sequence. It gave very high levels of mutagenesis. Using the plaque morphologies of the mutated stocks as an indicator, it was estimated that 5–50% of the phage carried mutated T4 lysozyme genes. This made it possible to identify more mutants

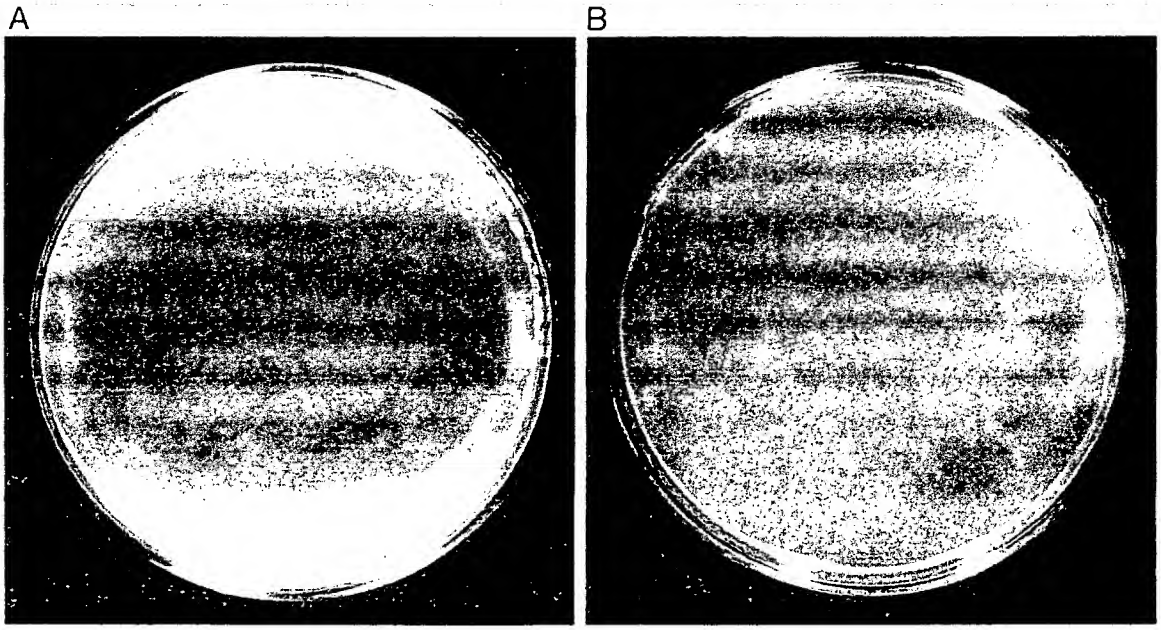



Fig. 1. Assay plates showing the discrimination between WT* and a thermostable mutant. **A**: WT*. **B**: Mutant N144E/WT*.

**Table 1.** *Activity and stability of mutant lysozymes*[a]

| Mutant | Activity (%) | Charge on mutant relative to reference protein | $\Delta T_m$, pH 5.4 (°C) | $\Delta T_m$, pH 6.78 (°C) | $\Delta H$, pH 5.4 (kcal/mol) | $\Delta S^0$, pH 5.4 (e.u.) | $\Delta\Delta G$, pH 5.4 (kcal/mol) |
|---|---|---|---|---|---|---|---|
| WT* | 100 | 0 | 0.0 | 0.0 | 134 | 397 | 0.0 |
| R14K | 106 | 0 | −0.08 | −0.53 | 135 | 401 | −0.03 |
| K16E (WT) | 142[b] | −2 | 1.1[c] | | | | |
| E22K | 40 | +2 | 1.37 | 1.58 | 141 | 400 | 0.57 |
| T26S | 75 | 0 | 1.35 | 1.35 | 143 | 406 | 0.57 |
| N40D | 124 | −1 | 1.14 | 1.28 | 132 | 378 | 0.44 |
| A41D | 105 | −1 | 0.71 | 1.10 | 138 | 398 | 0.29 |
| A41V (WT) | ND | 0 | 0.58 | ND | NC | NC | 0.26 |
| D70K (H31N) | 33 | +1 | 1.7[d] | ND | NC | NC | 0.36 |
| R80K | 44 | 0 | −0.43 | −0.34 | 136 | 408 | −0.17 |
| R80K/R119H | 104 | 0 | −1.20 | −1.25 | 133 | 406 | −0.47 |
| A93T | 105 | 0 | 0.13 | 0.16 | 138 | 408 | 0.06 |
| G113E | 165 | −1 | 0.79 | 1.03 | 126 | 364 | 0.30 |
| N116D (WT) | ND | −1 | 1.6[e] | | | | |
| R119H | 105 | 0 | −0.74 | −0.83 | 132 | 399 | −0.29 |
| K147E (WT) | 120[b] | −2 | −1.6[c] | | | | |
| T151S | 86 | 0 | 0.93 | 0.97 | 141 | 407 | 0.39 |
| F153L | ND | 0 | 0.88 | 0.51 | 137 | 393 | 0.35 |
| N163D | 193 | −1 | −0.50 | −0.49 | 141 | 424 | −0.21 |

[a] All mutants are in the WT* background except where "(WT)" is explicitly indicated in the first column. Activity was measured as the rate of hydrolysis of a suspension of *E. coli* cell walls (Tsugita et al., 1968). The charge on the mutant relative to the reference protein is the formal charge change obtained by counting changes in ionizable groups. The mutants identified in this study were most well characterized at pH 5.4 in 0.10 M NaCl, 0.010 M $H_{0.14}Na_{0.86}OAc$ (sodium chloride, sodium acetate buffer). This buffer is at or close to the pH value for which T4 lysozymes that have the His 31–Asp 70 salt bridge intact are most stable (Anderson et al., 1990). At this pH, all measurements (except for footnoted mutants) were the average of 4–8 independent determinations, with WT* having been measured 15 times. The melting temperature, $T_m$, of WT* in pH 5.4 buffer was 65.15 ± 0.2 °C and in pH 6.78 buffer (0.15 M KCl, 0.01 M $KPO_4$) was 62.19 ± 0.2 °C. The $\Delta T_m$ values were determined by subtraction of the $T_m$ of the reference protein from the $T_m$ of the mutant, and have estimated uncertainty of ±0.3 °C. $\Delta H$, the change in the enthalpy of unfolding at the melting temperature, and the melting temperature itself were both determined from van't Hoff analysis of the melting curves (Becktel & Baase, 1987; Dao-pin et al., 1990). $\Delta S^0$ is the isothermal change in the entropy of unfolding calculated at the melting temperature of WT*, assuming a constant value for $\Delta C_p$ of 3.5 kcal/mol-deg. $\Delta\Delta G$ is the isothermal difference in the change in the free energy of unfolding, $\Delta G$ (mutant) minus $\Delta G$ (reference), determined at the $T_m$ of the reference protein. Since $\Delta T_m$ values are small, the $\Delta\Delta G = \Delta T_m * \Delta S$ formula of Becktel and Schellman (1987) was used. Errors for $\Delta H$, $\Delta S^0$, and $\Delta\Delta G$ are estimated to be ±5 kcal/mol, ±15 e.u., and ±0.08 kcal/mol, respectively, from the averaging of the standard deviations of the means for sets of individual mutants. ND, not determined; NC, not comparable due to the change in background from WT* to WT or to H31N/D70N.

[b] Dao-pin et al. (1991b).

[c] 25 mM KCl, 20 mM $KPO_4$, pH 5.3 (Dao-pin et al., 1991b).

[d] Relative to the $T_m$ of the reference protein H31N/D70N (57.2 °C).

[e] Nicholson et al. (1991).

while screening fewer phage. It also made it possible to return to the WT* background, and so avoid the obvious complication arising from revertants associated with the H31N/D70N background.

Three different sets of stocks were made with this method, each using a different primer located either upstream from, or within, the T4 coding sequence. By adjusting the conditions of the initial primer extension reaction, it was possible to localize the sites of misincorporation to areas within about 200 bases of the coding sequence downstream from each primer. Approximately 11,000 plaques from stocks made from the three different primers were screened, of which 150 were selected for rescreening. From these, 66 were picked for sequencing. About a third of these were mutants and consisted of one or more isolates of 16 different mutants. These are listed in Table 1 and shown in Figure 3 and Kinemage 1. The rest were pseudo-wild type.

### *Stability and activity*

It was found that five of the mutants isolated (K16E, A41V, N116D, K147E, F153L) had previously been constructed by site-directed methods, although usually in the background of WT rather than WT*. Four of these have been shown to be more stable than wild type (Table 1). These were not studied further.

The other 13 variants were transferred into the expression vector pHSE5 (Muchmore et al., 1989) and trans-

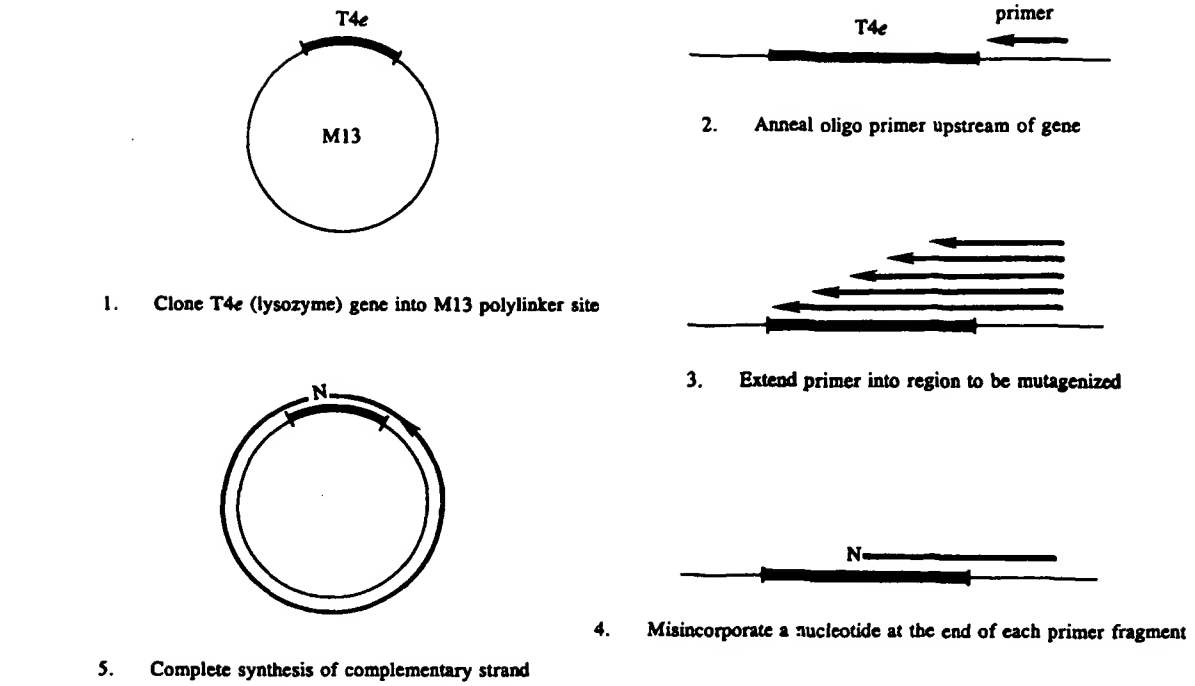



**Fig. 2.** Misincorporation mutagenesis scheme. The basic scheme is the same for both the thionucleotide and AMV reverse transcriptase protocols, with the major difference being the polymerase used for the misincorporation reaction. In both cases, uracil-containing M13 template is used to eliminate replication of the wild type-containing copy of the parent phage genome.

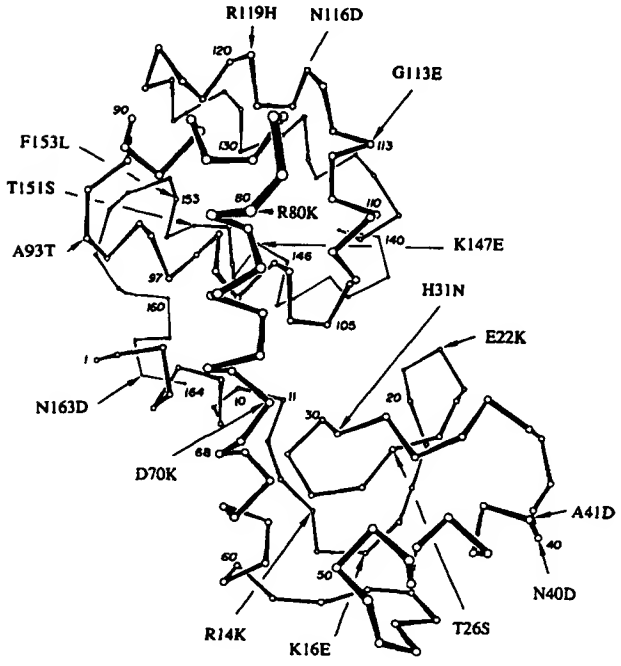



**Fig. 3.** Backbone of phage T4 lysozyme, showing the locations of all mutants discussed in the text.

formed into *Escherichia coli* strain RR1, and the mutant proteins were purified (Poteete et al., 1991). Thermal stabilities were determined by measuring the thermal denaturation of the proteins under reversible conditions by monitoring the change in circular dichroism at 223 nm as a function of temperature (Dao-pin et al., 1990; Eriksson et al., 1993). Activities were determined by the method of Tsugita et al. (1968). The results are summarized in Table 1.

Ignoring H31N/D70K, of the 17 mutants selected, 12 are bona fide thermostable variants, with melting temperatures 0.8–1.4 °C higher than wild type (the estimated error in $\Delta T_m$ is ±0.3 °C; Table 1). The other seven typically have either approximately the same stability as WT* or are somewhat less stable, with K147E ($\Delta T_m$ = −1.6 °C) being least stable. The single mutants R80K and R119H, as well as the double mutant R80K/R119H, were each identified as independent isolates. All are slightly less stable than WT*, with $\Delta T_m$ for the double mutant (−1.2 °C) essentially equaling the sum of the two singles (−0.4 °C and −0.7 °C) (Table 1). Activities are all within a factor of two of WT*.

The structures of three representative mutants are described in the companion paper (Pjura & Matthews, 1993).

## Discussion

### *M13 mutagenesis/screening system*

The overall objective was to develop a convenient system for the identification of thermostable mutants of T4 lysozyme. The most difficult requirement for this system is sensitivity. Except for variants that introduce disulfide bridges (Matsumura et al., 1989), the most thermostable mutant of T4 lysozyme that has been characterized to date has a melting temperature 2.8 °C above WT* at pH 5.4 (Anderson et al., 1993). Most thermostable variants are within 1–2 °C of wild type (e.g., Bell et al., 1990; Matthews, 1993). The assay therefore was required to detect stability differences of as little as 1 °C in order to identify typical thermostable mutants. This mitigated against the use of a selection protocol that relied upon large differences in activity or stability and required the use of a screening method that allowed the detection of more subtle differences.

Although substantial experience has suggested that the halo assay for lysozyme activity is a useful guide, it is not easy to discriminate between small differences in stability. In the present case the screen using the plate assay was the most difficult part of the protocol to develop because it has a number of variables that are difficult to control. The halo size can be quite variable, even for the plating of a single mutant stock. Conditions such as the particular batch of plates, the temperature of the incubator, the length of time after the plates are chloroformed before they are examined, and other factors can affect average halo size. The most troublesome difficulty, however, is the limited range of temperatures within which the phage and bacteria will grow. M13 phage will not grow below 32 °C, and the bacteria will not grow above 42 °C. This limits the usefulness of the assay to this range of temperature. This is a particular problem when WT* is used as background, because the temperature at which its halo size starts to decrease, which is the optimum temperature for discriminating between differences in stability, appears to be at or perhaps slightly above 42 °C. It was in part for this reason that an attempt was made to use the alternative H31N/D70N background. This mutant has a melting temperature of about 5 °C below that of WT* and allows good discrimination at an incubation temperature of 37 °C. However, selection in a background of H31N/D70N has obvious disadvantages, including frequent selection of revertants such as H31N/D70K. On balance, it was preferable to use WT* and to accept the slightly lower level of discrimination that it provided.

There is also an important difference between the M13-based and T4-based assays (Alber & Wozniak, 1985) in regard to the point in time at which the lysozyme begins to diffuse into the agar plate. Bacteriophage T4 is lytic, so each round of virus production is accompanied by the release of lysozyme. This lysozyme can diffuse outward, but the surrounding bacterial lawn has not yet been exposed to chloroform and so is resistant to attack by the enzyme. In the M13-based technique described here, however, the addition of chloroform synchronizes the release of lysozyme to a relatively well-defined time point. This means that different phenotypes tend to be distinguished by changes in halo diameter rather than in the degree of clarity of the halo.

### *Optimization of mutagenesis*

An objective was to find a method of mutagenesis that was quick and easy to use, was efficient, would be generally applicable over the entire gene, and would give predominantly single-site mutants.

Of the methods tested, the most useful was a variation on the Klenow thionucleotide misincorporation method using AMV reverse transcriptase for the primer extension and misincorporation reactions. It provided most of the mutants isolated. One disadvantage is that it is prone to making more than one mutation in the gene. In one case this resulted in the isolation of the double mutant R80K/R119H. Another problem is that mutations occasionally can also occur outside of the intended area of the M13 genome targeted by the initial primer-extension reaction. In a number of cases where putative mutants were found to code for WT* lysozyme, the phage genome upstream from the T4 coding region was also sequenced. In all such cases, mutations were found in the *lac* operator gene and were presumably the cause of the observed overexpression of the protein. In several other cases tested, where mutants were identified in the coding sequence, no mutations were found in the upstream region.

Each of the methods tested would be expected to generate single-base-change mutants, which limits the range of amino acid substitutions possible at a given codon.

### *Mutants isolated*

In total, 18 candidates for thermostable mutants were isolated (Fig. 3; Kinemage 1). Of these, five had previously been constructed by site-directed mutagenesis. Because four out of five of these site-directed mutants were constructed in different backgrounds, it obviates any possible concern that they might have been picked up by contamination. This also gave confidence that the overall approach was effective. Of the 18 candidates, 11 (K16E, E22K, T26S, N40D, A41D, A41V, H31N/D70K, G113E, N116D, T151S, F153L) are in fact thermostable mutants, 4 (R14K, R80K, A93T, N163D) are essentially equivalent to the background mutant, and 3 (R80K/R119H, R119H, K147E) are slightly (0.8–1.2 °C) less stable. As shown in Figure 3, the mutants come from a number of contexts in the protein: 11 (R14, K16, N40, R80, A93, G113, N116, R119, R80/R119, K147, N163) are surface-exposed residues (≥60% of side chain exposed), 4 (E22, A41, D70,

T151) are partially buried (20–60% exposed), and 2 (T26, F153) are core-packing residues (≤20% exposed); 5 come from the N-terminal domain (residues 1–59), 7 from the C-terminal domain (residues 81–164), and 1 from the long helix joining the two (residues 60–80); 2 are acidic residues (E22, D70), 6 are basic residues (R14, K16, R80, R80/K119, R119, K147), 5 are neutral polar residues (T26, N40, N116, T151, N163), and 4 are hydrophobic (A41, A93, G113, F153).

A number of false positives were found using the assay, and these fall into one of two categories: either candidates that were subsequently found to code for the wild-type (or WT*) gene, or actual mutants that had melting temperatures the same as or slightly less than wild type. The first group may in part be due to inadvertent mutations created outside of the T4e region leading to increased expression of the T4 gene.

The second group of false positives probably result from a number of different factors. The first is the creation of surface mutations that alter activity and/or the mobility of the protein through the agar substrate. Dao-pin et al. (1991b) observed that mutants of T4 lysozyme that included Lys → Glu or Arg → Glu substitutions had apparent activities ranging from 90 to 170% of wild type when measured with the cell wall turbidity assay of Tsugita et al. (1968). When measured on agar–peptidoglycan plates using an in vitro halo assay (Becktel & Baase, 1985) related to the one used here, however, the apparent activities increased dramatically, ranging from 270 to 13,600% that of wild type. It appeared that the mutants reduced the net positive charge on T4 lysozyme and thereby enhanced its mobility. Thus, a halo of larger than usual size on a lysozyme assay plate might occur for several reasons. The lysozyme might have enhanced stability, it might have enhanced activity, or it might simply have enhanced mobility through the gel. Table 1 shows evidence for a correlation between charge and activity. Six of the mutants cause a reduction in positive charge, and all of these are slightly more active than wild type (105–193%). This is very comparable with the observations of Dao-pin et al. (1991b). Only two of the mutants cause an increase in positive charge, and both lead to a reduction in activity (40% and 33% of WT).

Of the seven false positives (R14K, R80K, R80K/R119H, A93T, R119H, K147E, N163D; Table 1), all but one cause changes in the distribution of surface charge. K147E has been shown by Dao-pin et al. (1991b) to be substantially more mobile in agarose than WT, which might also explain why it appeared as a false positive. N163D also appears to be more active than WT by about a factor of two. The other false positives do not decrease positive charge; neither are they more active than wild type. Possibly the redistribution of surface charge affects mobility on the assay plate, but this is speculation.

Another factor in the generation of false positives is the limitation of the assay. As noted, the combination of phage and host strain limits the maximum temperature of incubation to 42 °C. This is a problem for WT* because the temperature at which its halo size begins to decrease, which is the most sensitive for detecting changes in melting temperature, is above this value. As a result, the discrimination between WT* and the more stable mutants is relatively poor. For mutant N144E, which is 1.5 °C more stable than WT*, the typical increase in halo diameter is from 7 mm to 8 or 9 mm. This is very close to the average variation in halo size. If a mutant lysozyme with lower melting temperatures is used as the "parent" or "reference," the discrimination is much greater. Mutant H31N/D70N is 1.7 °C less stable at pH 5.4 than H31N/D70K, but the halos differ in diameter by a factor of two (2.5 mm vs. 5 mm). This is in fact part of the reason why this mutant was explored as the "reference" protein (ultimately WT* was preferred because it is the standard reference, even though it allows poorer discrimination). Yet another possible factor in the generation of false positives is that the conditions used for the plate assay are not the same as those used for the CD measurements (see below).

All this aside, however, a relatively simple procedure has been developed to generate a library of mutants and to assay them directly. Using this system, approximately 120 candidates were selected. All that had mutations within the lysozyme gene were found to have melting temperatures within 1.5 °C of WT*.

The entropies of unfolding of two of the mutants, namely N163D and G113E, stand out relative to WT* (Table 1). For N163D, $\Delta S^0$ is about two standard deviations above WT*, whereas for G113E it is about two standard deviations below. In the crystal structure of WT*, residues 162–164 at the C-terminus of the molecule are very mobile (Weaver & Matthews, 1987). In mutant N163D, however, the introduced aspartate apparently interacts favorably with the dipole of the C-terminal helix, and, as a result, in the crystal structure of this mutant residues 162–164 are better ordered than in WT* (data not shown). This decrease in entropy of the folded mutant structure could explain the observed increase in the entropy of unfolding. For G113E, the likely effect of the mutant is on the unfolded structure (Matthews et al., 1987). Substitution of Glu for Gly would be expected to decrease the entropy of the unfolded state, which in turn could explain the observed decrease in the entropy of unfolding for this variant.

Finally, what is the physical basis of the assay? The observed initial increase in halo size with increasing temperature, followed by a sudden decrease in halo size above a certain temperature, is consistent with the behavior that one would expect of an enzyme at or near a critical temperature. The melting temperature, $T_m$, is measured for the purified enzyme, in buffer, at a defined pH. In contrast, in the plate assay the enzyme is in the presence of substrate and a host of other factors, and is at a temper-

ature 20 °C or so below $T_m$. The plate assay also extends over 12–24 h, very much longer than the refolding rate of the protein. It is therefore likely that the plate assay is sensitive not only to changes in $T_m$ but also to kinetic and other factors as well.

## Materials and methods

### *Reference proteins and construction of M13 vector system*

The principal reference protein used in this work was the cysteine-free double mutant C54T/C97A, referred to as WT* (Matsumura & Matthews, 1989). This eliminates the chemically reactive cysteines present in wild-type lysozyme but has little effect on structure or stability. Another reference or background strain was obtained by introducing two additional mutants, H31N and D70N, and is referred to as H31N/D70N. In wild-type lysozyme the salt bridge between His 31 and Asp 70 contributes 3–5 kcal/mol to the stability of the protein (see Kinemage 1; Anderson et al., 1990). The fragments were cloned into the polylinker region of the single-strand phage M13mp18 between the BamHI and HindIII restriction sites, using standard methods (Maniatis et al., 1982).

### *M13 lysis plate assay*

Phage stocks were assayed as follows. Typically, 100 μL of a stock, diluted to a titer of 250–1,000 pfu/mL with either Low TE buffer (10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA) or UV buffer (21.1 mM $Na_2HPO_4$, 11.1 mM $KH_2PO_4$, 68.4 mM NaCl, 28.7 mM $K_2SO_4$, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.01% gelatin, pH 7.0), was combined with 200 μL of a JM101 overnight culture, grown in YT broth (5 g NaCl, 8 g tryptone, 5 g yeast extract per liter, pH 6.8–7.0) or 2× YT broth (5 g NaCl, 16 g tryptone, 10 g yeast extract per liter, pH 6.8–7.0), and 5 mL of top YT agar (5 g NaCl, 8 g tryptone, 5 g yeast extract, 7 g agar per liter, pH 6.8–7.0), at approximately 55 °C, and plated onto 3.5-inch glass YT agar plates (5 g NaCl, 8 g tryptone, 5 g yeast extract, 15 g agar per liter, pH 6.8–7.0). The plates were allowed to set for 30 min at room temperature, then were inverted and incubated in a dry plate incubator at time(s) and temperature(s) as described. After incubation, the plates were allowed to cool for 20–30 min at room temperature. Finally, 1 mL of $CHCl_3$ was added to the lid of each, and the plates were incubated overnight at room temperature. Representative examples are shown in Figure 1.

### *Hydroxylamine mutagenesis*

Mutagenesis reactions were prepared by combining 10 μL of an M13 phage stock, 65 μL of 0.2 M $Na_3PO_4$, pH 6.0, and 25 μL of 1 M hydroxylamine-HCl, pH 6.0. The reaction mixture was incubated at 37 °C in a water bath. At zero time and at successive intervals, 10-μL aliquots were taken and diluted 10,000-fold in 50% UV buffer/50% glycerol and stored at −20 °C. Stocks were assayed for titer by serial dilution and plating onto lawns of JM101 (Messing, 1983), as for the lysis plate assay.

### *Thionucleotide misincorporation mutagenesis*

The procedure is adapted from that of Abarzua and Marians (1984) and is summarized in Figure 2. Because it proved not to be very effective in the present case, details will not be given.

### *Reverse transcriptase misincorporation mutagenesis*

The procedure followed that of Lehtovaara et al. (1988), with modifications. First, an annealing reaction was prepared containing 3.3 μL of 1.5 pmol/μL kinased mutagenic primer, 2.8 μL of 0.15 pmol/μL U* M13 template, 2.5 μL of 10× Klenow buffer (0.5 M Tris-HCl, pH 7.5, 0.1 M $MgCl_2$, 10 mM dithiothreitol, 1 mg/mL bovine serum albumin), and 16.4 μL sterile $H_2O$ in a 1.5-μL Eppendorf tube. The reaction was heated to 65–75 °C in a water bath and allowed to cool to room temperature over approximately 0.5–2.5 h.

Next, the annealed primer/template mix was split and used in two primer-extension reactions, each containing 12.5 μL of the primer/template mix, 3.75 μL of 10× Klenow buffer, either 5 or 12 μL of 2 μM each dATP, dCTP, dGTP, and dTTP, and water to make 49 μL total volume. Primer extension was carried out by adding 1 μL of 1 U/μL Klenow enzyme to each and incubating for 5 min at room temperature. The reactions were stopped by adding 2 μL of 0.5 M NaEDTA, pH 7.0, to each. The reactions were each extracted with 1:1 phenol/chloroform, the two supernatants combined, and the pooled reactions extracted with chloroform. The recovered supernatant was then ethanol precipitated overnight at −20 °C. Finally, the primer/template reaction was recovered by spinning at 4 °C in a microfuge, washed in 80% ice-cold ethanol, and dried in a Speedvac. The pellet was resuspended in 80 μL of Low TE buffer.

The mixture was then divided and used in four misincorporation reactions, each containing 20 μL of the extended primer/template mix, 2.5 μL of 10× AMV buffer (0.5 M Tris-HCl, pH 8.3, 60 mM $MgCl_2$, 0.6 M KCl, 10 mM dithiothreitol, 0.9 mg/mL bovine serum albumin), and 2.5 μL of 2 mM dNTP (one type only per reaction). Misincorporation was performed by adding 1 μL of 7 U/μL AMV reverse transcriptase (Promega) to each. The reactions were incubated for 90 min in a 37 °C water bath, 0.5 μL of reverse transcriptase was added to each, and the reactions were continued for another 90 min. Finally, the templates were extended beyond the misincorporation sites ("chased") by adding, to each, 2.5 μL of

5 mM each of all four dNTPs and incubating for another 20 min at 37 °C.

Next, double-strand synthesis and ligation reactions were performed on each of the four misincorporation mixes by combining each with 2.5 μL of 10× AMV buffer, 5 μL of 10 mM ATP, and 13 μL of sterile $H_2O$. The reactions were carried out by adding, to each, 0.5 μL of 1 U/μL Klenow enzyme (Boehringer Mannheim) and 0.5 μL of 2 U/μL T4 DNA ligase (Boehringer Mannheim). The reactions were carried out overnight at 15 °C in a circulating water bath. Finally, the reactions were stopped by adding 2 μL of 0.5 M NaEDTA, pH 7.0, to each, and the mixes were stored at −20 °C.

Finally, the four misincorporation reaction mixes were used to transform competent cells and to raise phage libraries for subsequent screening. Competent JM101 cells were made using the $CaCl_2$ method, as previously described. The Ca-treated cells were kept in ice water for approximately 90 min before use. Each of the separate dNTP misincorporation reactions was added to a 2-mL portion of the competent cells and incubated for 30 min in ice water. The cells were heat shocked by transferring to a 37 °C water bath for 7 min, then returned to room temperature. A fresh 4-mL volume of 2× YT broth was added to each and incubated for 6 h at 37 °C on a rocker table. Finally, phage stocks were made from the incubations by spinning down 1 mL of each for 5 min at room temperature in a microfuge and combining 0.5 mL of each supernatant with 0.5 mL of glycerol and storing at −20 °C.

In a variation on the above, the initial primer extension reaction was performed using AMV reverse transcriptase instead of Klenow as follows: following annealing as described above, the mixture was used in the two primer extension reactions as before, except that 10× AMV buffer was used in place of Klenow buffer, 1 μL of 7 U/μL reverse transcriptase was used in place of Klenow, and the reaction was incubated for 10 min instead of 5 min. After the incubation, the misincorporation reaction was performed directly, without the extraction and precipitation steps, by combining for each misincorporation reaction 12.5 μL of each molarity extension reaction with 2.5 μL of 2 mM of one dNTP. They were incubated for 90 min in a 37 °C water bath, an additional 0.5 μL of reverse transcriptase was added to each, and the incubations continued for another 90 min. The reactions were chased for 20 min at 37 °C by adding, to each, 2.5 μL of a solution 5 mM in each of the four dNTPs. Finally, 10× AMV buffer and 10 mM ATP were added as above with water to make 49 μL total volume, and the mixtures were incubated overnight with Klenow enzyme and T4 ligase as before.

## Acknowledgments

We are most grateful to Joan Wozniak, Sheila Pepiot, and Joel Lindstrom for their excellent technical assistance and Drs. Hale Nicholson and Brian Shoichet for helpful advice. This work was supported in part by a PHS fellowship (GM11836) to P.P., by a grant from the NIH (GM21967) to B.W.M., and by the Lucille P. Markey Charitable Trust.

## References

Abarzua, P. & Marians, K.J. (1984). Enzymatic techniques for the isolation of random single-base substitutions in vitro at high frequency. *Proc. Natl. Acad. Sci. USA 81*, 2030–2034.

Alber, T. & Matthews, B.W. (1987). Structure and thermal stability of phage T4 lysozyme. *Methods Enzymol. 154*, 511–533.

Alber, T. & Wozniak, J.A. (1985). A genetic screen for mutations that increase the thermal stability of phage T4 lysozyme. *Proc. Natl. Acad. Sci. USA 82*, 747–750.

Anderson, D.E., Becktel, W.J., & Dahlquist, F.W. (1990). pH-induced denaturation of proteins: A single salt bridge contributes 3–5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry 29*, 2403–2408.

Anderson, D.E., Hurley, J.H., Nicholson, H., Baase, W.A., & Matthews, B.W. (1993). Hydrophobic core repacking and aromatic-aromatic interaction in the thermostable mutant of T4 lysozyme Ser 117 → Phe. *Protein Sci. 2*, 1274–1284.

Becktel, W.J. & Baase, W.A. (1985). A lysoplate assay for *Escherichia coli* cell wall-active enzymes. *Anal. Biochem. 150*, 258–263.

Becktel, W.J. & Baase, W.A. (1987). Thermal denaturation of bacteriophage T4 lysozyme at neutral pH. *Biopolymers 26*, 619–623.

Becktel, W.J. & Schellman, J.A. (1987). Protein stability curves. *Biopolymers 26*, 1859–1877.

Bell, J.A., Brennan, R.G., Dao-pin, S., Faber, R., Karpusas, M., Matsumura, M., Nicholson, H., Pjura, P.E., Roderick, S.L., Tronrud, D.E., Weaver, L.H., Wilson, K., Wozniak, J.A., Zhang, X-J., & Matthews, B.W. (1990). Structural and genetic analysis of protein-protein and protein-DNA interactions. In *Frontiers in Drug Research, Alfred Benzon Symposium 28* (Jensen, B., Jorgensen, F.S., & Kofod, H., Eds.), pp. 281–294. Munksgaard, Copenhagen.

Bryan, P.N., Rollence, M.L., Pantoliano, M.W., Wood, J., Finzel, B.C., Gilliland, G.L., Howard, A., & Poulos, T.L. (1986). Proteases of enhanced stability: Characterization of a thermostable variant of subtilisin. *Proteins Struct. Funct. Genet. 1*, 326–334.

Dao-pin, S., Baase, W.A., & Matthews, B.W. (1990). A mutant T4 lysozyme (Val 131 → Ala) designed to increase thermostability by the reduction of strain within an α-helix. *Proteins Struct. Funct. Genet. 7*, 198–204.

Dao-pin, S., Sauer, U., Nicholson, H., & Matthews, B.W. (1991a). Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis. *Biochemistry 30*, 7142–7153.

Dao-pin, S., Soderlind, E., Baase, W.A., Wozniak, J.A., Sauer, U., & Matthews, B.W. (1991b). Cumulative site-directed charge-change replacements in T4 lysozyme suggest that long-range electrostatic interactions contribute little to protein stability. *J. Mol. Biol. 221*, 873–887.

Eriksson, A.E., Baase, W.A., & Matthews, B.W. (1993). Similar hydrophobic replacements of Leu 99 and Phe 153 within the core of T4 lysozyme have different structural and thermodynamic consequences. *J. Mol. Biol. 229*, 747–769.

Lehtovaara, P.M., Koivula, A.K., Bamford, J., & Knowles, J.K.C. (1988). A new method for random mutagenesis of complete genes: Enzymatic generation of mutant libraries *in vitro*. *Protein Eng. 2*, 63–68.

Liao, H., McKenzie, T., & Hageman, R. (1986). Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. *Proc. Natl. Acad. Sci. USA 83*, 576–580.

Makino, Y., Negoro, S., Urabe, I., & Okada, H. (1989). Stability-increasing mutants of glucose dehydrogenase from *Bacillus megaterium* IWG3. *J. Biol. Chem. 264*, 6381–6385.

Maniatis, T., Fritsch, E.F., & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Matsumura, M. & Aiba, S. (1985). Screening for thermostable mutant of a kanamycin nucleotidyltransferase by the use of a transforma-

tion system for a thermophile, *Bacillus stearothermophilus*. *J. Biol. Chem. 260*, 15298–15303.

Matsumura, M., Becktel, W.J., Levitt, M., & Matthews, B.W. (1989). Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc. Natl. Acad. Sci. USA 86*, 6562–6566.

Matsumura, M. & Matthews, B.W. (1989). Control of enzyme activity by an engineered disulfide bond. *Science 243*, 792–794.

Matthews, B.W. (1993). Structural and genetic analysis of protein stability. *Annu. Rev. Biochem. 62*, 139–160.

Matthews, B.W., Nicholson, H., & Becktel, W.J. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl. Acad. Sci. USA 84*, 6663–6667.

Messing, J. (1983). New M13 vectors for cloning. *Methods Enzymol. 101*, 20–79.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E., & Dahlquist, F.W. (1989). Expression and $^{15}N$ labelling of proteins for proton and nitrogen-15 NMR. *Methods Enzymol. 177*, 44–73.

Nicholson, H., Anderson, D.E., Dao-pin, S., & Matthews, B.W. (1991). Analysis of the interaction between charged side-chains and the $\alpha$-helix dipole using designed thermostable mutants of phage T4 lysozyme. *Biochemistry 30*, 9816–9828.

Pjura, P. & Matthews, B.W. (1993). Structures of randomly generated mutants of T4 lysozyme show that protein stability can be enhanced by relaxation of strain and by improved hydrogen bonding via bound solvent. *Protein Sci.* 2, 2225–2231.

Poteete, A.R., Dao-pin, S., Nicholson, H., & Matthews, B.W. (1991). Second-site revertants of an inactive T4 lysozyme mutant restore activity structuring the active site cleft. *Biochemistry 30*, 1425–1432.

Streisinger, G., Mukai, F., Dreyer, W.J., Miller, B., & Horiuchi, S. (1961). Mutations affecting the lysozyme of phage T4. *Cold Spring Harbor Symp. Quant. Biol. 26*, 25–30.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., & Inouye, M. (1966). Frameshift mutations and the genetic code. *Cold Spring Harbor Symp. Quant. Biol. 31*, 77–84.

Tsugita, A., Inouye, M., Terzaghi, E., & Streisinger, G. (1968). Purification of bacteriophage T4 lysozyme. *J. Biol. Chem. 243*, 391–397.

Weaver, L.H. & Matthews, B.W. (1987). Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. *J. Mol. Biol. 193*, 189–199.

MOLECULAR AND CELLULAR BIOLOGY, Mar. 1995, p. 1499–1512
0270-7306/95/$04.00+0

Vol. 15, No. 3

# Genetic Dissection of Thyroid Hormone Receptor β: Identification of Mutations That Separate Hormone Binding and Transcriptional Activation

RAVINDRA UPPALURI AND HOWARD C. TOWLE*

*Department of Biochemistry and Institute of Human Genetics, University of Minnesota, Minneapolis, Minnesota 55455*

Received 4 November 1994/Returned for modification 7 December 1994/Accepted 15 December 1994

**The thyroid hormone receptors (TR) are members of the nuclear receptor family of ligand-mediated transcription factors. The large region of TR that lies C-terminal to its DNA-binding domain subserves functions of ligand binding, dimerization, and transactivation. Little is known regarding the structural or functional determinants of these processes. We have utilized genetic screening in the yeast *Saccharomyces cerevisiae* to identify residues involved in these functions. Random mutations of the rat TRβ1 isoform between amino acid residues 179 and 456 were screened, and mutants with reduced hormone-dependent activation of reporter gene activity were isolated. In this paper we describe the characterization of a class of mutants that exhibit a dissociation between hormone binding and transcriptional activation. These mutants retained hormone binding (>15% of the wild-type level) yet failed to transactivate a reporter gene. A number of these mutations occurred within the D region, which links the DNA-binding and ligand-binding domains of the receptor. One subset of these mutations abrogated DNA binding, supporting a role of the D region in this process. The remainder retain DNA binding and thus highlight residues critical for receptor activation. In addition, an unexpected group of "superactivator" mutations that led to enhanced hormone-dependent activation in *S. cerevisiae* were found. These mutations localized to the carboxy-terminal portion of the receptor in a region which contains elements conserved across the superfamily of nuclear receptors. The hormone-dependent phenotype of these superactivator mutations suggests an important role of this segment in ligand-mediated transcriptional activation.**

Thyroid hormone receptors (TR) are members of the nuclear receptor family of ligand-mediated transcription factors (17). TR interact with DNA recognition elements known as thyroid hormone response elements (TRE) to regulate transcription of specific genes and thereby influence diverse aspects of development and homeostasis. TR may function as either homodimers or heterodimers with the retinoid X receptor (RXR) in binding to TRE sequences. By analogy to the steroid receptors (30), the TR sequence has been divided into five regions designated A/B (rat TRβ1 [TRβ] residues 1 to 101), C (102 to 169), D (170 to 237), and E (238 to 456) (throughout this paper, we have used the numbering system of Murray et al. [43] for amino acid positions). The C segment comprises the DNA-binding domain, a 68-amino-acid region containing two zinc finger elements that are highly conserved among family members. This domain confers DNA-binding specificity on various receptors in the family (40, 63). Other functions of TR, including ligand binding, dimerization, and transcriptional activation, have been attributed to the poorly defined D and E regions (33). However, attempts to identify the critical amino acid residues responsible for these various functions have met with limited success, suggesting considerable overlap in these domains. Our goal in initiating these studies was to define mutations of TR that would distinguish residues important for hormone binding and transcriptional activation.

Attempts to define domains of TR important for its various functions have been based on mutagenesis experiments that created deletions or amino acid substitutions within the coding sequences. Deletion of the A/B region of TRβ had no effect on its activity in cotransfection experiments, suggesting that this region was not critical to the functioning of the receptor (61). On the basis of a region of homology between TR and the steroid receptors, residues 281 to 300 of the TRβ isoform were found to be important for heterodimerization with RXR (45, 46). Further information has been gained from analyses of mutant receptors from individuals afflicted with the syndrome of generalized resistance to thyroid hormone (GRTH). In this autosomal dominant genetic disorder, TRβ gene mutations result in the production of receptors deficient in their responses to thyroid hormone. GRTH mutations highlight two regions of human TRβ—amino acids 310 to 347 and 417 to 453 (42, 67). The resulting receptors behave as dominant repressors of TR-specific gene transcription.

Limitations in the analysis of in vitro and GRTH mutations may have precluded the identification of important residues of TR. For studies using site-directed mutagenesis, the time-consuming nature of testing each mutation necessarily limits the number of mutations that can be evaluated. GRTH mutations are limited by the size of the patient population and the requirement that these mutations give a dominant negative phenotype. Thus, we sought to establish a scheme whereby the entire D and E regions could be investigated in an unbiased and expedient fashion.

The expression and functional analysis of various mammalian transcription factors in the yeast *Saccharomyces cerevisiae* have revealed a basic conservation in the processes of transcription and transcriptional activation in eucaryotes. This genetically malleable system has proven useful for studying the mechanism of action of steroid receptors. For example, Yamamoto and colleagues defined important residues in the DNA- and ligand-binding domains of the glucocorticoid recep-

* Corresponding author. Mailing address: 4-225 Millard Hall, Department of Biochemistry, University of Minnesota, Minneapolis, MN 55455. Phone: (612) 625-3662. Fax: (612) 625-2163.

tor (19, 56). Vegeto et al. have described an altered-specificity mutant of the progesterone receptor that activates transcription in response to the antagonist RU486 (65). Wilson et al. (69) have described the binding site for the NGFI-B factor and have identified amino acids important for this function. Ince et al. (26) and Wrenn and Katzenellenbogen (70) have used yeast phenotypic screens to identify dominant negative estrogen receptors and have highlighted residues of the estrogen receptor important for hormone binding. All of these studies have been possible only in light of the fact that genetic and phenotypic screens can be performed in *S. cerevisiae* with relative ease.

Human TRβ and rat TRβ have also been shown to function in *S. cerevisiae* (47, 49). We describe in this paper a yeast system for the functional characterization of rat TRβ. We performed a mutational analysis of the D and E regions of TRβ to identify mutations that led to defects in transcriptional activation. In particular, we have identified mutations that retain DNA and hormone binding but are deficient in activation. These mutations identify several specific regions of TR that are critical to its function as a ligand-dependent transcriptional factor.

## MATERIALS AND METHODS

**Plasmids.** All subcloning and transformations were done essentially as described by Sambrook et al. (54). Constructions of individual plasmids were as follows.

(i) **pG2M.** pG2M is a modified version of the yeast expression vector pG2 described by Schena et al. (57). *Bam*HI-digested pG2 was blunt ended by using the Klenow fragment of DNA polymerase I, and an *Mlu*I site was created by insertion of an *Mlu*I linker. This vector contains the yeast glycerol-3-phosphate dehydrogenase promoter to drive expression of exogenous DNA and a *TRP1* gene for selection.

(ii) **pG2M/TRβ1.** Rat TRβ1 sequences were amplified by PCR with four primers, designated 1 to 4 (see Fig. 2). Primer 1 contains an *Mlu*I site, primer 4 contains a *Sal*I site, and the overlapping primers 2 and 3 contain *Bgl*II sites. In separate reactions, regions spanned by primers 1 and 2 or by primers 3 and 4 were amplified with Vent DNA polymerase (New England Biolabs). Subsequently, these products were digested with the requisite enzymes (either *Mlu*I and *Bgl*II or *Bgl*II and *Sal*I). These fragments were then ligated into an *Mlu*I-*Sal*I-digested pG2M vector.

(iii) **TRβ truncations.** Carboxyl-terminal truncations were performed to create receptors of 200, 320, and 410 amino acids. The TRβ200 and TRβ320 truncations were created by amplification with primer 3 and appropriately positioned 3′ primers. The TRβ410 truncation was created by amplifying a clone that had a termination codon inserted at amino acid 410. All of these fragments were cut with *Bgl*II and *Sal*I and inserted into *Bgl*II-*Sal*I-cut pG2M/TRβ1.

(iv) **GAL4-TRβ fusions.** GAL4-TRβ fusion constructs were prepared in a three-step fashion. First, a *Bam*HI-*Cla*I fragment, encompassing the coding region for GAL4 amino acids 1 to 147, was isolated and cloned into *Bam*HI-*Cla*I-digested plasmid pSP73 (Promega). This construct was then digested with *Acc*I, which cuts in the polylinker of pSP73 and at position −54 relative to the GAL4 initiation codon. An *Mlu*I site was then introduced at position −54 via linker addition. Finally, an *Mlu*I-*Bgl*II fragment from this construct, encompassing the GAL4 coding region for amino acids 1 to 147, was exchanged into the corresponding region in pG2M/TRβ1. These manipulations result in the fusion of GAL4 residues 1 to 147 with TRβ residues 172 to 456. Three amino acids (Met-Ile-Ser) were inserted at the junction of the fusion. This same strategy was used to create fusions with specific TRβ mutants.

(v) **pTRE-cyc1/lacZ.** The yeast reporter vector, which contained two copies of the $TRE_{pal}$ element upstream of the yeast *cyc1* promoter, was obtained from M. Privalsky, University of California, Davis (49).

(vi) **pLGSD5.** The construct to test the function of GAL4 fusions contained the upstream activating sequence of the *GAL1/GAL10* gene located 5′ to a *CYC1* promoter/β-galactosidase cassette (23).

(vii) **CMVS4.** CMVS4 is a derivative of the mammalian expression vector CMV4 (2). Since the *Sal*I site located in the polylinker is not unique in the plasmid, CMV4 was partially digested with *Sal*I, and the singly digested linear fragment was isolated. This DNA was blunt ended with Klenow DNA polymerase and resealed. Plasmids with a unique *Sal*I site in the polylinker were then identified by restriction enzyme mapping.

(viii) **CMVS4/TRβ.** TRβ and all mutants were introduced into CMVS4 as *Mlu*I-*Sal*I fragments.

(ix) **4XTREpal/TKCAT.** 4XTREpal/TKCAT contains four copies of the $TRE_{pal}$ element upstream of the thymidine kinase promoter, which drives expression of the bacterial chloramphenicol acetyltransferase coding region (39).

**PCR. (i) Normal PCR.** Reactions were performed according to the manufacturer's directions with Pfu polymerase (Stratagene) or Vent polymerase (New England Biolabs). Denaturation was at 94°C for 1 min, annealing was at 45°C for 1 min, and extension was at 72°C for 2 min, for 25 cycles.

**(ii) Mutagenic PCR.** The technique used for mutagenic PCR was a minor modification of that described by Leung et al. (38). Reaction mixtures contained 1 ng of template to be mutagenized, 16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8.8), 6.1 mM $MgCl_2$, 0.5 mM $MnCl_2$, 6.7 μM EDTA, 170 μg of bovine serum albumin per ml, 10 mM β-mercaptoethanol, primers at 1 μM, 1 mM each dGTP, dTTP, and dCTP, 400 μM dATP, and 2.5 U of *Taq* DNA polymerase (Cetus). Cycling was performed as for normal PCR.

**Creation of mutant library.** Rat TRβ1 sequences were amplified with primers 3 and 4 by using mutagenic PCR conditions. Mutant fragments were isolated by electrophoresis on a 1% agarose gel and cleaved with *Bgl*II and *Sal*I. These fragments were then inserted into *Bgl*II-*Sal*I-digested pG2M/TRβ1. Ligations were transformed into SURE competent *E. coli* (Stratagene). A pool of all transformants (>2,000 individual colonies) was created, and plasmid DNA was isolated.

**Yeast strains and media.** Strain SSL204 (*MATα his3 leu2 trp1 ade2 ura3*) was used throughout this work (1). For testing of GAL4 fusions, YM706 (*MATα ura3 his3 ade2 lys2 trp1 tyr1 gal4*) was used. This strain has a deletion of all GAL4 coding sequences. Methods for growth and manipulation of yeast strains were as described previously (3). Yeast transformations were performed by electroporation by the method of Becker and Guarente (8).

**Generation of antibody to rat TRβ.** The receptor was overexpressed in *Escherichia coli* [BL21(DE3)/pLysS] by using a T7 RNA polymerase system. Cells were disrupted, and insoluble material was collected by centrifugation. This material was subjected to sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis, and the receptor was isolated by electroelution. The preparation of polyclonal TRβ antibodies was performed by Babco (Houston, Tex.). Antibodies were then affinity purified from serum as described by Cama et al. (12).

**SDS-polyacrylamide gel electrophoresis and Western blotting (immunoblotting).** Gels were run according to the method of Laemmli (32). Two sources of extract were used. Either 400 μl of a yeast culture grown to an $A_{600}$ of 1 or 12.5 μg of whole-cell extract (see below) was loaded per lane. Proteins were transferred to nitrocellulose filters (Costar), and immunoblotting was performed with anti-rat TRβ antibodies at a dilution of 1:750 and the ECL detection system (Amersham).

**Preparation of yeast extracts.** Extracts were prepared as described by Olesen et al. (48) and were used for electrophoretic mobility shift assays (EMSA) and immunoblotting analysis. Yeast cells were grown in 100 ml of minimal medium to an $A_{600}$ of 1. Cells were harvested and washed in 5 ml of EB [0.2 M Tris-HCl (pH 8), 400 mM $(NH_4)_2SO_4$, 10 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, and 7 mM β-mercaptoethanol). The pellets were transferred to microcentrifuge tubes in 1 ml of EB and centrifuged to collect cells. After resuspension in 200 μl of EB with 1 mM phenylmethylsulfonyl fluoride, 1 μg of leupeptin per ml, and 1 μg of pepstatin per ml, a one-half volume of glass beads was added. The suspension was frozen in a dry-ice–ethanol bath and thawed on ice. Cells were then vortexed at 4°C for 20 min. A further 100 μl of EB was added, and cells were left on ice for 30 min. The suspension was centrifuged for 5 min, and the supernatant was transferred to a new tube and centrifuged for 1 h in a microcentrifuge. The supernatant was then made to 40% with $(NH_4)_2SO_4$ and gently rocked for 30 min. After a 10-min centrifugation, the pellet was resuspended in 300 μl of 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 8.0)–5 mM EDTA–7 mM β-mercaptoethanol–1 mM phenylmethylsulfonyl fluoride–1 μg of leupeptin per ml–1 μg of pepstatin per ml–20% glycerol. Dialysis was then performed against the same buffer, and aliquots were stored at −70°C.

**Analysis of $T_3$ binding.** A 5-ml culture of *S. cerevisiae* was grown to an $A_{600}$ of 1. Cells were pelleted, washed with 0.5 ml of ice-cold BB (20 mM Tris-HCl [pH 7.5], 50 mM NaCl, 2 mM EDTA, 5 mM β-mercaptoethanol, 10% glycerol, and 1 mM $MgCl_2$) and transferred to a microcentrifuge tube. After centrifugation, the pellets were resuspended in 100 μl of BB containing 0.4 M KCl and an equal volume of glass beads. The samples were then frozen at −70°C and thawed on ice prior to disruption. Phenylmethylsulfonyl fluoride was added to 1 mM, and the tubes were vortexed for 5 min at 4°C. Fifty microliters of BB was added, and samples were vortexed for 10 s and centrifuged for 5 min. The hormone binding assays were performed with 25 μl of extract and 0.5 nM $^{125}$I-labeled 3,5,3′-triiodothyronine ($^{125}$I-$T_3$) (Dupont/NEN) in a volume of 250 μl, and assay mixtures were incubated for 16 h at 4°C. Nonspecific binding was assessed by including a 1,000-fold excess of unlabeled $T_3$. Bound counts were determined by the nitrocellulose filter binding assay of Inoue et al. (27). All assays were performed in duplicate on extracts of yeast colonies from two different transformations. Results are expressed relative to that for wild-type TR, which is set at 100%.

**β-Galactosidase assays.** In situ assays were performed on yeast colonies that had been replica plated and grown on minimal plates in the absence or presence of 1 μM 3,5,3′-triiodothyroacetic acid (Triac) at 25°C for 3 to 4 days. Ten milliliters of a molten solution containing 0.5 M potassium phosphate (pH 7), 0.1% SDS, 2% dimethylformamide, 0.2% X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), and 0.5% agarose was poured over the plate. After the agarose hardened, the plates were incubated at 37°C for 4 to 5 h and then stored overnight at 4°C. Yeast colonies with decreased β-galactosidase activity were

easily distinguished. Quantitative assays were done as described by Himmelfarb et al. (25). All assays were repeated at least twice on yeast colonies from separate transformations. β-Galactosidase units were calculated as $(1{,}000 \times A_{420})/(\text{minutes} \times \text{milligrams of protein})$, and duplicates varied by <20% in this assay. Since the GAL4-TRβ fusions were expressed at a lower level than wild-type TRβ, Galacto-Light (Tropix Inc., Bedford, Mass.), a more sensitive luminescent substrate for β-galactosidase, was used. Yeast extracts were prepared in 100 mM potassium phosphate (pH 7.8)–0.2% Triton X-100–1 mM dithiothreitol and were assayed according to the manufacturer's directions in a Berthold luminometer. Activity was defined as follows: units = [luminescence units/(minutes × milligrams of protein)]/100.

**Recovery of pG2M/TRβ1 plasmids from *S. cerevisiae*.** Selected yeast colonies were streaked on plates containing 5.7 mM 5-fluoroorotic acid and grown for 3 days. Plasmid DNA was isolated by the spheroplast lysis method of Boeke et al. (10). Digestion with *Hin*dIII was performed to assess yields and to ensure that no rearrangements of the plasmid occurred in *S. cerevisiae*.

**DNA sequencing.** Sequencing was performed by the Sanger dideoxynucleotide method as adapted for modified T7 DNA polymerase (U.S. Biochemicals).

**Separation of multiple mutations.** In certain cases, multiple mutations that were present within a single mutant receptor were separated by swapping restriction fragments with the normal TRβ vector. *Bst*XI cuts at amino acid 280 in TRβ and once in pG2M. Consequently, plasmids were digested with *Bst*XI, and the appropriate fragments were exchanged with wild-type *Bst*XI fragment. Clones were sequenced to ensure the presence of the mutation. Mutations that were separated are designated with a lowercase letter (a or b) following the number of the mutant.

**EMSA.** Binding of wild-type and mutant TR in yeast extracts (2.5 μg) to a single-copy $TRE_{pal}$ probe was assessed by using the binding and electrophoresis conditions described by Bodenner et al. (9), except with 2 μg of poly(dI · dC) and without hormone in the binding buffer.

**Cell culture and transient transfection.** CV-1 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 50 U of penicillin per ml, and 50 μg of streptomycin per ml. On the day prior to transfection, $10^6$ cells were plated per 60-mm-diameter plate in medium supplemented with 100 nM dexamethasone and 10% fetal bovine serum stripped of endogenous thyroid hormones (55). Cells were transfected via calcium phosphate coprecipitation as described by Zilz et al. (73), except precipitate was left on cells for 12 to 14 h and dimethyl sulfoxide shock was not performed. Each plate received 3 μg of 4XTREpal/TKCAT, 0.5 μg of Rous sarcoma virus long terminal repeat-driven luciferase internal control plasmid, 50 ng of receptor expression vector, and 0.95 μg of CMVS4. Cells were then washed, and medium without or with 50 nM $T_3$ was added. After a 24-h culture, cells were harvested in Reporter Lysis Buffer (Promega, Madison, Wis.) according to the manufacturer's directions. Chloramphenicol acetyltransferase (CAT) assays were then performed as described previously (28), and results are expressed normalized to luciferase activity.

## RESULTS

**Rat TRβ function in *S. cerevisiae*.** In order to achieve expression and transcriptional competence of rat TRβ in *S. cerevisiae*, we utilized a system described by Privalsky et al. (49). In this system yeast cells are cotransformed with two plasmids. The first plasmid contains the coding sequences of TRβ driven by a constitutive promoter. The second plasmid has a basal yeast (*CYC1*) promoter upstream of the *E. coli* β-galactosidase gene. Two copies of the palindromic TRE (22) are located upstream of the basal promoter. In this fashion, expression of the β-galactosidase reporter gene was coupled to binding of functional TR to its DNA recognition site.

To demonstrate expression of the rat TR in *S. cerevisiae*, we measured $T_3$ binding in an extract from cells transformed with the TRβ expression plasmid. Scatchard analysis showed a binding affinity for $T_3$ indistinguishable from that demonstrated in mammalian cells (data not shown). In addition, extracts were analyzed by Western blotting. A rabbit polyclonal antibody for TRβ recognized a 50-kDa band in extracts from *S. cerevisiae* transformed with the TR expression vector (Fig. 1B). No bands were observed in extracts from *S. cerevisiae* containing the expression vector without TR coding sequences (see Fig. 4A). Functional activity was assessed by monitoring reporter gene expression in the absence and presence of thyroid hormone. As observed in *S. cerevisiae* by others (47, 49), the expression of the reporter gene was elevated in a receptor- and TRE-dependent fashion by rat TRβ (Fig. 1A). This activity contrasts to the

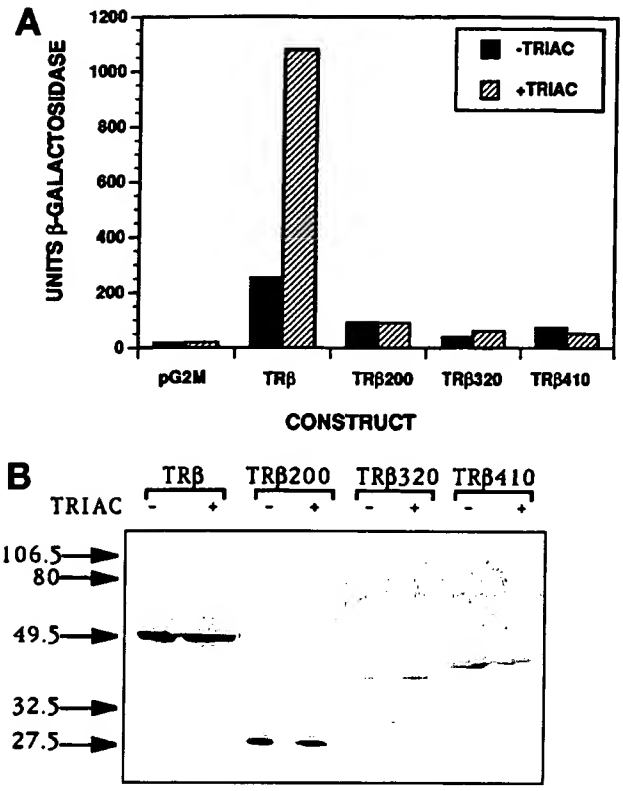



FIG. 1. (A) Activation of the $TRE_{pal}$ reporter by wild-type TR and various C-terminally truncated forms. Truncated TR were created to contain the first 200, 320, or 410 amino acids of TRβ. Expression vectors containing these constructs were cotransformed with the $TRE_{pal}$ reporter plasmid into *S. cerevisiae*. Transformants were grown without or with 1 μM Triac to an $A_{600}$ of 1 and assayed for β-galactosidase activity. pG2M represents activity from the $TRE_{pal}$ reporter in the absence of TRβ coding sequences. The β-galactosidase activity is expressed as units per milligram of protein and represent the average of duplicates that varied by <20%. (B) Western blot analysis of extracts from *S. cerevisiae* containing wild-type and truncated versions of TRβ. Yeast cells from the same cultures grown for the β-galactosidase assay were harvested by a brief centrifugation. The cell pellet was resuspended directly in sample buffer, boiled, and loaded on an SDS–10% polyacrylamide gel. After electrophoresis and transfer to nitrocellulose, the filter was probed with a polyclonal TRβ antibody and bands were visualized with the ECL detection system. The approximate molecular weights (in thousands) of prestained markers (Bio-Rad Laboratories) are indicated on the left.

observation that in mammalian cells unliganded receptor often causes an inhibition of basal transcriptional activity (11, 15). To test the hormone-dependent activity of TR in *S. cerevisiae*, we used the acetic acid analog of $T_3$, Triac, since it permeates *S. cerevisiae* more effectively than $T_3$ (49). In the presence of Triac, TRβ acts as a ligand-dependent transcriptional activator, giving a three- to fivefold increase in reporter gene activity.

To confirm that the D and E regions of TR are necessary for ligand-mediated transcriptional activation, we prepared a series of C-terminal truncations of TR and expressed these in *S. cerevisiae*. The truncations retain the first 200, 320, or 410 amino acids of rat TRβ. Expression of these constructs was confirmed by Western blotting (Fig. 1B). The TRβ truncation containing residues 1 to 200 is expressed at levels comparable to those of the wild-type TR, whereas the other two truncated TR are present at reduced levels. As shown in Fig. 1A, all of the truncated TR forms retain some hormone-independent activity, albeit reduced compared with that of the full-length

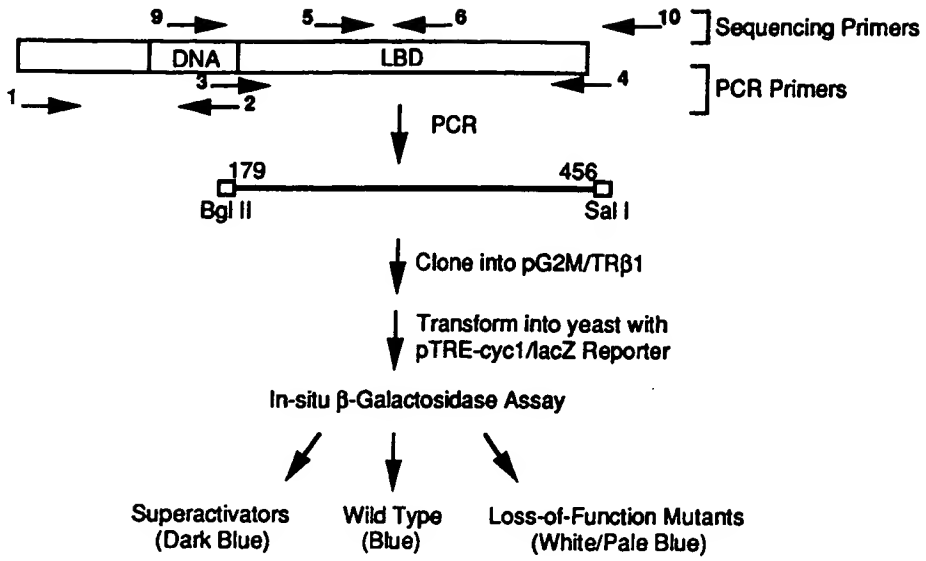



FIG. 2. Strategy for generating and screening TRβ mutations in the extended (D/E) ligand-binding domain. Primers 3 and 4 were used under mutagenic PCR conditions to amplify nucleotides 510 to 1383 of TRβ1. After digestion with *Bgl*II and *Sal*I, amplification products were cloned into *Bgl*II-*Sal*I-digested pG2M/TRβ. A library of mutants was created by passaging the ligated DNA through *E. coli*, pooling the cells, and isolating plasmid DNA. Yeast cells were then transformed with the library and the $TRE_{pal}$ reporter. Transformants were replica plated without and with 1 μM Triac. After growth for 3 days, an in situ β-galactosidase assay was performed. The phenotypes observed and subsequent classification are as indicated.

receptor. This observation suggests that these proteins retain some ability to interact with $TRE_{pal}$. Thus, a portion of the constitutive activity seen in *S. cerevisiae* is encoded by the A/B or C domain of TRβ. Upon addition of hormone, the truncated receptors did not stimulate reporter gene activity. Thus, as expected, all hormone-dependent activity is abrogated by truncations of the C-terminal region of TRβ.

**Design of the genetic screen.** Our primary interest in pursuing a yeast system for TR action was to use the genetic capabilities of *S. cerevisiae* to define regions of the receptor critical to its function. Specifically, we were interested in defining important amino acids located in the D and E regions, comprising the 277 amino acid residues downstream of the DNA-binding domain. To this end, we devised a mutant screening strategy which is summarized in Fig. 2. The basic features of this yeast screen have been described by Garabedian and Yamamoto (19).

In the first stage of this screen, we created a cassette version of the TRβ cDNA clone that contained a *Bgl*II site at nucleotides 513 to 518 and a *Sal*I site adjacent to the termination codon. These modifications did not alter the amino acid sequence of TRβ, and this clone was functionally indistinguishable from the original TRβ when expressed in *S. cerevisiae* (data not shown). Primers that flanked amino acids 179 and 456 were then used in a mutagenic PCR amplification reaction under conditions described by Leung et al. (38). Conditions were optimized to achieve a mutation rate of two or three nucleotides for the 831-bp region (see Materials and Methods). The PCR products were digested with *Bgl*II and *Sal*I and reintroduced into the yeast expression vector in place of the wild-type sequences. Passage through *E. coli* resulted in a library of mutants containing >2,000 individual TR expression plasmids. These mutant TR sequences were then cotransformed with the reporter plasmid containing $TRE_{pal}$ sequences into *S. cerevisiae*.

Transformed yeast colonies were isolated and replicated to plates without and with 1 μM Triac. This concentration of hormone is higher than that required for saturation of the receptor and thus eliminated the identification of mutants that had minor reductions in ligand-binding affinity. An in situ β-galactosidase assay was used to screen approximately 600 colonies. Since each plasmid had an average of two amino acid substitutions, roughly 1,200 mutations were present in the pool of mutants screened. Thus, each amino acid in the 277-residue region should be mutated, on average, four times in the colonies screened. Putative mutants (164 colonies) were either white or light blue relative to a wild-type control. Fortuitously, we identified a set of 18 colonies that appeared to give higher β-galactosidase activities than normal TRβ (see below).

Several additional screening steps were performed to eliminate mutations that were not of interest to us. First, Western blotting allowed us to identify and eliminate unstable or truncated versions of TRβ. Next, the TRβ expression plasmids were isolated and separately retransformed into *S. cerevisiae* along with the $TRE_{pal}$ reporter plasmid. Extracts from the transformed yeast cells grown with or without hormone were prepared and assayed for β-galactosidase activity by a quantitative in vitro technique. This step ensured that the deficient phenotype was due to mutation of the TR expression vector. This assay also helped to eliminate a small number of wild-type receptors which scored as false positives in the in situ assay. On the basis of the β-galactosidase assay, mutant receptors were divided into two groups: ones that retained some level of hormone-independent activity (70 mutants) and ones that had lost all hormone-independent activity (44 mutants). Hormone-independent activity requires nuclear localization and DNA binding. Since we were most interested in identifying residues that affected the transactivation function of TR, only mutants that retained hormone-independent activity were further characterized.

We next prepared cell extracts from the mutants that retained hormone-independent activity to assess their $T_3$-binding abilities. This assay divided the mutants into two major classes of approximately equal size: those that retained $T_3$ binding and

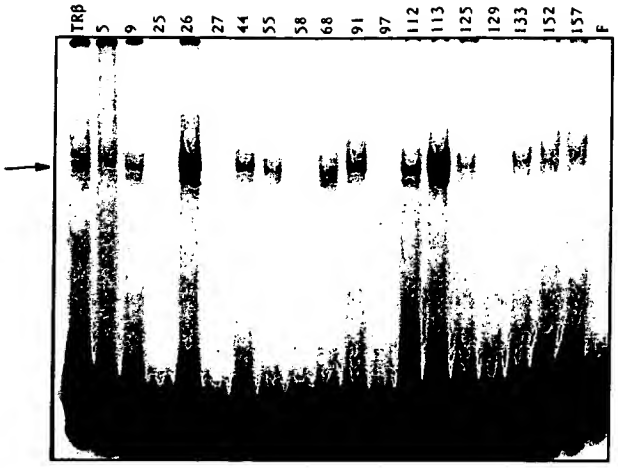



FIG. 3. EMSA of TRβ mutants with the $TRE_{pal}$ probe. Whole-cell extracts (25 μg) containing the receptor mutants indicated above the lanes were assessed for DNA binding by using the $TRE_{pal}$ sequence. Lane F (free) contains probe without any protein. The arrow indicates the position of the homodimer complex formed with the $TRE_{pal}$ sequence.

those that did not. We arbitrarily chose a binding activity of >15% of that of the wild-type receptor as the criterion for inclusion in the $T_3$-binding class. The focus of this paper is on 18 mutants that fell into this class and on the properties of 14 mutants which gave transactivation higher than that of the normal receptor. The characterization of the remaining mutants with alterations in hormone-binding properties will be discussed elsewhere (64).

**Analysis of DNA binding of activation-deficient, $T_3$-binding-positive mutants.** We next sought to determine whether the defects in activation in the mutants that retained $T_3$ binding were due to an inability to bind to the DNA response element. DNA binding was assessed with the $TRE_{pal}$ oligonucleotide in the EMSA. No bands were observed with an extract prepared from *S. cerevisiae* lacking TR coding sequences (data not shown). Extracts from *S. cerevisiae* bearing the wild-type TR gave a single complex that could be inhibited by unlabeled $TRE_{pal}$ oligonucleotide but not a control oligonucleotide. The migration of this band was coincident with that observed with $TRE_{tap}$ as a probe (data not shown). TRβ has been shown to bind preferentially as a homodimer to this latter sequence (5, 71). Thus, we conclude that the wild-type TRβ band observed with the $TRE_{pal}$ sequence represents a homodimer.

Since mutagenesis was targeted to regions outside the DNA-binding domain of TR, we did not expect to find mutations that directly affected this function. To corroborate this assumption, extracts from *S. cerevisiae* with various mutant TR were prepared and tested for binding to the $TRE_{pal}$ probe. Most of the mutants (13 of 18) bound to an extent similar to that of the wild-type receptor (Fig. 3). Minor changes in the mobilities of certain mutant forms can be seen and may represent alterations in the conformation or charge of these proteins. These mutants, being normal in both hormone binding and DNA binding, represent alterations that affect the transactivation function of TR. Surprisingly, 3 of the 18 mutants (25a, 27a, and 58a) did not detectably bind $TRE_{pal}$. Two other mutants (97 and 129a) bound weakly to the labeled probe. For these five mutants, the defect in binding to $TRE_{pal}$ can account for the deficiency in hormone-dependent transactivation. Thus, there are at least two subsets of mutant TR: those that bind and those that do not bind $TRE_{pal}$.

**Sequence analysis of mutant TR.** To examine the locations of mutations giving rise to mutant phenotypes, the receptor sequences were determined by using four primers that allowed complete sequencing of the 831-bp region subjected to mutagenesis (Fig. 2). Many of the mutant TR contained more than one mutation. Several of these multiple mutations were separated, and each mutation was tested individually. In every case, the defective phenotype segregated with a single amino acid change.

The distribution of these mutations was nonrandom. Eleven of the 18 mutants contain mutations in the D region between amino acids 179 and 206. This "hinge" region, denoted such since it connects the DNA-binding and ligand-binding domains (30), contains a high concentration of positively and negatively charged amino acids. Fourteen of 24 residues in this segment are charged. Six of the mutations found in this region resulted in a conversion from a charged to a neutral amino acid residue. Eight of the mutations did not affect the net charge of the receptor, and one change (N-190 to D [N190D]) resulted in a region with more charge. Thus, mutations were not preferentially found at charged residues. Amino acids changed in more than one mutant include K-184, N-190, and L-199. Table 1 summarizes the mutations, levels of hormone binding, and reporter gene activities for these 11 D region mutants.

Interestingly, all of the mutants defective in DNA binding had mutations that mapped to the hinge region. For example, two different mutations that abrogated DNA binding were found at residue N-190 (25a and 27a). This region of the receptor has recently been suggested to regulate the rotational flexibility between the C and E domains of TR that is important for DNA binding (31). The remaining six mutants with mutations in the hinge region bound normally to the $TRE_{pal}$ probe. In addition, all of these hinge region mutants bound hormone comparably to the wild-type receptor. Indeed, Scatchard analysis of mutant 9 demonstrated that its $T_3$-binding affinity was unchanged relative to that of wild-type TRβ (data not shown). Thus, the D region plays a dual role in DNA binding and transactivation of TR.

The properties of the seven remaining mutants which fail to transactivate are summarized in Table 2. These mutations are not clustered but are dispersed widely throughout the E region. Mutant 5, which has a P448S mutation, is coincident with a mutant from the GRTH syndrome. The properties of the P448H mutant have been described by Nagaya and Jameson (44). Mutant 113 contains two substitutions in a region corresponding to inactivating mutations introduced into rat TRα, as described by Lee and Mahdavi (36). Mutants 152 and 157 contain mutations that are 6 amino acids apart and immediately downstream from the heterodimerization domain defined by O'Donnell et al. (45, 46). The locations of changes in mutants 26, 112, and 125 have not been implicated in previous work. However, mutants 26 and 68 both contain changes of L-261, in mutant 26 to a proline and in mutant 68 to a glutamine.

The 18 mutants exhibited a spectrum of hormone-independent and hormone-dependent activities. Several of the mutants (e.g., mutants 26, 55, 97, 133, 152, and 157) retained a modest degree of response to hormone but much less than that observed with the wild-type receptor. Others showed no hormonal induction of reporter gene activity. A range of activities in the absence of hormone was also noted. Many receptors showed a hormone-independent activity comparable to that of the wild-type receptor, indicating that the defect was limited to the hormone-activated response. Others showed a decrease in hormone-independent activity as well as a depression in the hormone-stimulated level. Given these differences, a compar-

TABLE 1. TR mutants with mutations within the D (hinge) region[a]

| TR | β-Galactosidase (U) | | % T3 binding | Mutation(s) |
|---|---|---|---|---|
| | −TH | +TH | | |
| TRβ (wild type) | 353 | 1,141 | 100 | |
| Mutants that do not bind $TRE_{pal}$ | | | | |
| 25a | 50 | 55 | 46 | N190D |
| 27a | 100 | 177 | 114 | N190S |
| 58a | 250 | 333 | 110 | K179R, R184G |
| 97 | 104 | 256 | 119 | K206R |
| 129a | 108 | 161 | 102 | R194G, L199P |
| Mutants that bind $TRE_{pal}$ | | | | |
| 9 | 91 | 104 | 100 | I203T |
| 44a | 163 | 130 | 124 | L199P |
| 55a | 206 | 515 | 167 | I187T |
| 68 | 374 | 418 | 93 | E192G, E198G, L261Q |
| 91 | 125 | 198 | 119 | R184G, E215G |
| 133 | 170 | 467 | 139 | R191H |

[a] Wild-type or mutant TR expression vectors were introduced into *S. cerevisiae* together with the $TRE_{pal}$ reporter vector and grown in the absence (−TH) or presence (+TH) of 1 μM Triac. β-Galactosidase activity in extracts of these yeast cells was measured. Hormone binding was estimated by incubating extracts in the presence of 0.5 nM $^{125}I$-$T_3$ for 16 h and then separating bound hormone from free hormone by filtration on nitrocelluose. All values for T3 binding are expressed relative to that for wild-type TRβ, which was set at 100%. All measurements were repeated with yeast cultures from a separate transformation and gave comparable results. β-Galactosidase activity in yeast colonies that were transformed with the reporter vector and the parental TR expression vector was 15 to 20 U.

ison of mutant TR activities as fold inductions has little meaning. Some of the variation in reporter gene activity may reflect modest differences in levels of mutant receptors expressed in *S. cerevisiae*. However, as can be seen from Fig. 4A, most of the mutant forms were expressed at levels similar to that of the wild-type receptor. Thus, the differences in activities of these mutants likely reflect changes in their transcriptional activation potentials.

**TR mutants that superactivate.** Although our screen was designed for isolating loss-of-function mutants, we serendipitously recovered a set of TR mutants with enhanced transactivation ability. These colonies appeared dark blue relative to colonies with wild-type TR in the in situ screen. These mutant forms have been designated superactivators. Because these mutants may modify the transactivation function of the receptor, we have pursued the characterization of this unexpected group of mutants. The hormone binding, reporter gene activities, and mutational changes of these receptors are summarized in Table 3.

Most of these superactivators have increased transcriptional potencies only in the presence of hormone. However, some show increased hormone-independent activity as well. Only one, mutant S20, does not show any hormone-dependent increase in reporter gene activity. Although some mutants demonstrate only modest increases in hormone-dependent activation, mutants S4 and S15 have strikingly increased hormone-dependent activities. These mutants show a 10-fold increase in activation in response to hormone. Interestingly, both contain the same K419E change. Mutant S3 also contains this mutation, but it activates only three- to fourfold more than wild-type TR. Presumably the impact of the K419 change is blunted by the presence of the two other changes present in this protein.

A possible explanation for the superactivator phenotype is that these mutants have increased levels of protein expression. However, this did not appear likely in light of the fact that the hormone-independent activity was unchanged in the majority of these TR. To formally exclude this possibility, we performed Western analysis on extracts of *S. cerevisiae* that contained these proteins (Fig. 4B). With the exception of mutant S17, all

TABLE 2. TR mutants with mutations within the E region[a]

| TR | β-Galactosidase (U) | | % T3 binding | Mutation(s) |
|---|---|---|---|---|
| | −TH | +TH | | |
| TRβ (wild type) | 353 | 1,141 | 100 | |
| 5 | 333 | 399 | 37 | P448S |
| 26a | 322 | 484 | 43 | L261P |
| 112 | 296 | 441 | 142 | N359S |
| 113 | 345 | 480 | 27 | Q236P, F240S, N326D S356T |
| 125 | 346 | 440 | 15 | T322A, F394L |
| 152 | 214 | 697 | 60 | L300P, A431V |
| 157 | 237 | 795 | 51 | A226V, E306G |

[a] See Table 1, footnote *a*.

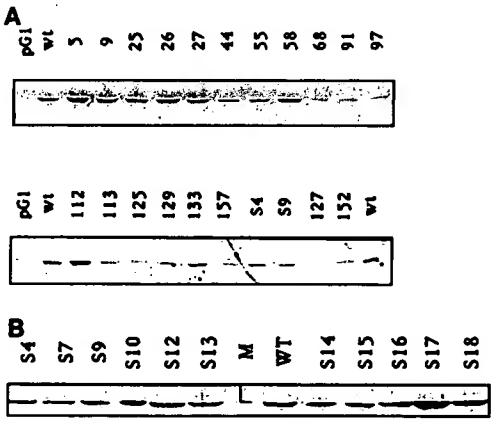



FIG. 4. (A) Western blot of TR mutants that lost wild-type activation function but retained $T_3$ binding. Yeast whole-cell extracts were prepared from cells containing the mutant TRs indicated above the lanes. pG1, extract from *S. cerevisiae* containing expression vector without TRβ coding sequences; wt, extract from *S. cerevisiae* containing TRβ. Each lane was loaded with 25 μg of extract, and a Western blot was performed as described in Materials and Methods. (B) Western blot of superactivator TRβ mutants. These extracts were prepared as described for Fig. 1B. M, markers; WT, wild type. Mutants S3, S19, and S20 are not included in this Western blot.

TABLE 3. TR mutants that superactivate transcription[a]

| TR | β-Galactosidase (U) | | % T3 binding | Mutation(s) |
|---|---|---|---|---|
| | −TH | +TH | | |
| TRβ (wild type) | 444 | 1,132 | 100 | |
| S3 | 206 | 4,384 | ND[b] | G246R, R277G, K419E |
| S4 | 631 | 12,149 | 42 | K419E |
| S7 | 80 | 2,039 | 21 | D211G, Y401C, E455G |
| S9 | 317 | 2,883 | 97 | E452A |
| S10 | 1,099 | 3,340 | 61 | K415E |
| S12 | 931 | 2,242 | 56 | K179T, I275T, E452G |
| S13 | 585 | 2,629 | 174 | F412S, R433H |
| S14 | 102 | 2,004 | 40 | E364G |
| S15 | 1,200 | 11,906 | 9 | K419E, F450S |
| S16 | 380 | 2,927 | 166 | E198G, R424Q |
| S17 | 585 | 3,682 | 196 | K218R, H436R |
| S18 | 515 | 2,079 | 172 | T421M |
| S19 | 212 | 2,169 | ND | F288L, C293R |
| S20 | 6,663 | 2,559 | ND | V279E, K283R, K301Q |

[a] See Table 1, footnote *a*.
[b] ND, not determined.

of these mutants are expressed at levels less than or equivalent to that of wild-type TR. In fact, the strongly active S4 is actually expressed at a level slightly less than that of the normal receptor.

The $T_3$-binding abilities of these receptors were also analyzed. Most of the mutant TR in this group bound $T_3$ comparably to the wild-type TR. Surprisingly, one of these mutants (S15) showed little $T_3$ binding, despite the fact that it gave a strong induction of reporter gene activity in the presence of hormone. We presumed that this mutant retained a reduced affinity for hormone. This was confirmed by analysis of reporter gene activities of mutants S4 and S15 and wild-type TR grown with various amounts of Triac. The hormone-dependent activation by S15 required a significantly higher concentration of Triac than that by wild-type TR or S4 (data not shown). Whereas wild-type TR and S4 responded well at 10 nM Triac, S15 displayed no activity at this hormone concentration.

The most intriguing feature of the superactivators is the distribution of the mutations that give rise to the phenotype. Eleven of the 14 superactivators have mutations between amino acid residues 415 and 456 at the C-terminal end of the receptor. This region in the steroid/thyroid receptor family of proteins has been implicated in transactivation (7, 16, 52).

**Effects of TRβ mutations on activities of GAL4-TR fusion proteins.** To verify that the mutations identified affected the transactivation function of TR, several mutants were tested in the context of a heterologous DNA-binding and dimerization domain. For this purpose, we chose the well-characterized DNA-binding and dimerization domain of the yeast GAL4 transcriptional activator. The domain spanning amino acids 1 to 147 of GAL4 has been shown to function when fused to a wide variety of transactivating proteins, including members of the steroid receptor family (66). In addition, GAL4-TR fusion proteins have been shown to exhibit both transactivation and repression functions when introduced into mammalian cells (6). In a GAL4-TRβ fusion, the GAL4 portion of the fusion should not require elements outside amino acids 1 to 147 for efficient DNA binding. This strategy thus allows us to directly test effects of mutations on transactivation.

A fusion protein was prepared by joining the coding sequences for GAL4 amino acids 1 to 147 to the region coding for amino acids 172 to 456 of TRβ in the yeast expression

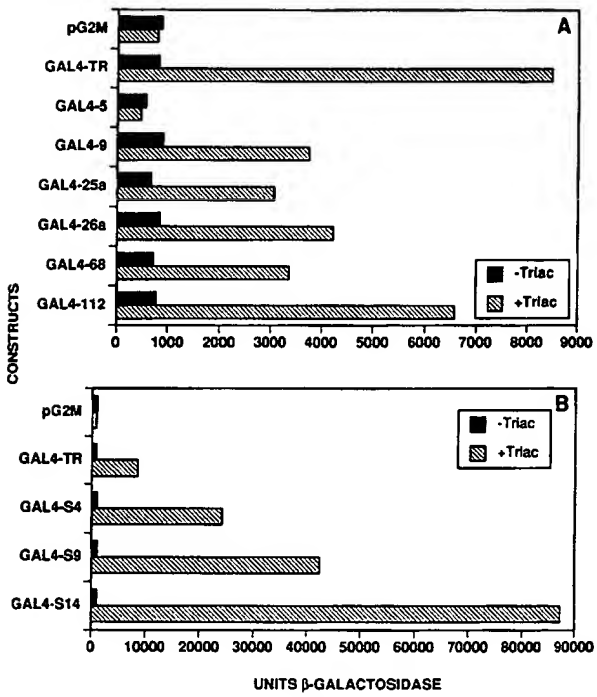



FIG. 5. Activities of GAL4-TRβ fusions. (A) $T_3$-binding, transactivation-deficient class. (B) Superactivator class. All constructs fused GAL4 amino acids 1 to 147 with TRβ (wild type or mutant) amino acids 172 to 456. The indicated constructs were cotransformed with pLGSD5 into the GAL4-deficient yeast strain YM706. Transformants were grown without or with 1 μM Triac to an $A_{600}$ of 1, and β-galactosidase activity was determined. pG2M represents activity of the pLGSD5 reporter in the absence of any fusion protein. Units are expressed per milligram of protein and are representative of two independent experiments.

vector. This fusion construct was introduced into *S. cerevisiae* together with a β-galactosidase reporter construct containing binding sites for GAL4. Since these fusions were expressed at a level lower than that of full-length TRβ (data not shown), a more sensitive assay for β-galactosidase was utilized (see Materials and Methods). The fusion protein with wild-type TRβ sequences was indeed found to function as a hormone-dependent activator in *S. cerevisiae* (Fig. 5A). The degree of hormone-dependent induction with this fusion protein was about 10-fold, which is somewhat more than that observed for the wild-type receptor. This difference may reflect the fact that the GAL4-TR fusion protein shows little hormone-independent activity, which suggests that this activity is a product of the A/B or C region of TRβ. These data emphasize that the C-terminal region of TRβ contains a hormone-dependent activation domain that can function in *S. cerevisiae* independent of the A/B and C regions.

Several loss-of-function TR mutants were subcloned into the GAL4-TR fusion vector and tested for their abilities to support hormone-dependent activation (Fig. 5A). All forms showed decreased activity relative to the wild-type receptor sequences, although the extent of the effect varied with different mutants. For example, the fusion containing mutant 5 sequences is devoid of activity, whereas the fusion with mutant 112 retains about 75% of the wild-type activity. Several individual mutations that occur in the D region of the receptor displayed between 30 and 50% of the activity of the fusion containing wild-type sequences. Because the 25a TR mutant was also defective in $TRE_{pal}$ binding, the reduced activity observed with

GAL4-25a suggests that this residue plays a role in both DNA binding and transactivation. Western blots of extracts from the fusion constructs indicated no significant differences in the levels of proteins expressed relative to that of the wild-type GAL4-TRβ fusion (data not shown). The diminished hormone induction supported by the mutant fusion proteins strongly implicates the D and E domains in the process of transcriptional activation by TRβ.

Several mutants from the class of superactivators were also tested as fusion proteins with GAL4 (Fig. 5B). All of these fusions displayed hormone-dependent activations that were greater than the activity of the wild-type fusion protein. Interestingly, the strongest superactivator in this context appeared to be S14 rather than S4. One explanation for the differences in activation between full-length TRβ mutants and fusion proteins of these mutants may be the absence of the A/B/C domains of TRβ. The fact that each superactivator behaves as such in the contexts of both the full-length product and the fusion construct again argues that the mutations specifically affect the transactivation potentials of these molecules.

**Mammalian cell activity of mutants.** To test whether TR mutations that led to defective transcriptional activation in *S. cerevisiae* effected a loss of function in mammalian cells, we tested the activities of the full-length mutant TR in CV-1 cells. This fibroblastic cell line does not express significant levels of TR. A CAT reporter plasmid containing four copies of the $TRE_{pal}$ element was introduced into CV-1 cells along with a set of TRβ mutants. As shown in Fig. 6A, this reporter is not activated in the presence of hormone when the expression vector CMVS4 alone is cotransfected. With TRβ, a strong hormone-dependent induction of CAT activity is observed. All mutants tested showed a reduced response to hormone. However, the magnitude of these effects did not correlate with those observed in *S. cerevisiae*. In general, the mutant TR were less affected in CV-1 cells than in *S. cerevisiae*. However, none of these constructs yielded more than half the induction in mammalian cells compared with the wild-type TR control. For example, mutant 5 had no induction in *S. cerevisiae* yet showed a seven- to eightfold induction in CV-1 cells. To eliminate the possibility that these data reflect unequal levels of the different receptors in mammalian cells, we tested the levels of mutants 5, 9, 26, and 112 expressed and found no major discrepancies relative to the wild-type receptor (data not shown).

We also tested several superactivator mutants (S4, S10, and S13) in CV-1 cells (Fig. 6B). Surprisingly, none of these mutants were superactivators in this context. Indeed, all effected a reduced level of activation in CV-1 cells. Mutant S4, the strongest hormone-dependent activator in *S. cerevisiae*, actually was the weakest activator in CV-1 cells. These data demonstrate that the alteration that results in a more powerful activation domain in *S. cerevisiae* results in an opposite phenotype in mammalian cells.

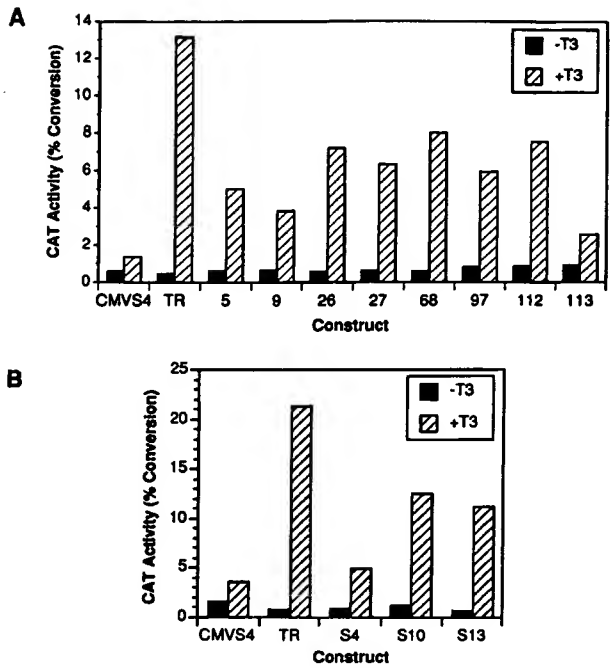



FIG. 6. (A) Activities of $T_3$-binding-competent, transactivation-deficient mutants in CV-1 cells. Wild-type TRβ and the indicated mutants were transferred to the CMVS4 expression vector and transfected into CV-1 cells along with the 4XTREpal/TKCAT reporter. Cells were grown in the absence or presence of 50 nM $T_3$, and CAT assays were performed. CAT activity was normalized to the amount of luciferase activity expressed from a Rous sarcoma virus-luciferase internal control plasmid. (B) Activities of some superactivator mutants in CV-1 cells. Superactivators S4, S10, and S13 were transferred to CMVS4, and transfections were performed as described for panel A.

## DISCUSSION

**Transcriptional activation by TR in *S. cerevisiae*.** The process of transcriptional activation involves the formation of protein-protein contacts between specific domains of the transcriptional activator and components of the transcriptional machinery. The specific domains of several transcriptional activators involved in this process have been identified (for a review, see reference 41). These domains are classified largely on the basis of the amino acid content (i.e., acidic, glutamine rich, or proline rich), since no structural information about these regions is presently available. For TR and many members of the steroid receptor family, little is known regarding the nature of the transactivation domain. In part, this lack of information stems from the intimate association between residues important for hormone binding and transcriptional activation. The extensive D and E regions of TR not only contain dispersed residues involved in hormone binding but also contain residues important for transactivation. Thus, attempts to dissociate hormone binding and transactivation have been largely unsuccessful. We undertook an unbiased screen for transactivation mutations of *S. cerevisiae*. Our mutant collection shows unambigiously that select changes in the D and E regions can affect transactivation without altering hormone binding.

An understanding of the meaning of mutant phenotypes generated in *S. cerevisiae* requires a discussion of the ways in which the behavior of TR in *S. cerevisiae* either mimics or does not mimic the function of TR in its normal setting. The basic features of TR activation are retained in *S. cerevisiae*. Induction of reporter gene activity is dependent on both TR expression and the presence of a $TRE_{pal}$ sequence. Similar responses to hormone are observed with reporter vectors containing two other TREs known to be active in mammalian cells: a direct repeat of the AGGTCA motif with a 4-bp spacer or an inverted repeat separated by 6 bp (data not shown). Furthermore, induction occurs in the presence of active thyromimetic analogs, including $T_3$, $T_4$, and Triac, but not in the presence of inactive analogs such as reverse $T_3$ (reference 49 and data not shown). Thus, the abilities of TR to bind to response elements and activate when bound to ligand are maintained in *S. cerevisiae*. Hormone-dependent activation in *S. cerevisiae* requires se-

quences from the C-terminal region of the receptor, as shown by the loss of activity in C-terminally truncated forms of TR. This conclusion was confirmed by the observation that the D/E region confers hormone-dependent transcriptional activation on a heterologous DNA-binding and dimerization domain. Again, this property has been found for TR activity in mammalian cells as well. Finally, several of the $T_3$-binding-competent, transactivation-defective mutants identified in the yeast system behave as mutant TR in mammalian cells. Thus, the process of hormone-activated transcription likely involves conserved determinants of the receptor in both yeast and mammalian cells.

A number of features, some of known origin and some of unknown origin, distinguish TR action in *S. cerevisiae* from TR action in mammalian cells. The most obvious difference is the transactivation that occurs in *S. cerevisiae* when hormone is not present. Recent data suggest that TR possesses a discrete repressing or silencing function in its D/E region (13, 14). This repressing activity can function on a heterologous promoter in transfected cells or in a cell-free transcription assay (18) and may be mediated by a direct interaction of TR and the basal transcription factor TFIIB (4). We suggest that the repressing activity of TR does not function in *S. cerevisiae*. This difference could represent the deficiency in *S. cerevisiae* of a soluble repressor protein that recognizes TR or a lack of evolutionary conservation between the repressing domain of TR and the basal yeast transcription apparatus. In the absence of this repressing function, the receptor which binds DNA in the unliganded state displays a low level of transcriptional activation. The nature of this activity remains undefined to date. A truncated form of TRβ containing amino acids 1 to 200 retained approximately 30% of the hormone-independent activity of the full-length receptor. Thus, at least a portion of the hormone-independent activity arises from the A/B or C region. Further evidence supporting this idea comes from experiments with GAL4 fusion proteins. The GAL4-TR chimeric protein, which lacks the A/B and C regions, was inactive in the absence of hormone. The TRα1 isoform has been reported to contain a hormone-independent N-terminal activation domain that functions only on a modified TRE sequence in mammalian cells (53). However, no such function has been reported for TRβ. Thompson and Evans (61) deleted the N terminus of human TRβ and showed activity equivalent to that of the wild-type receptor in CV-1 cells. Since some activation domains are cell specific, the presence of an activation region in the N terminus of TR that functions in certain mammalian cell types cannot be excluded.

Another difference between yeast and mammalian cells is the behavior of TR mutants. Although the transactivation-deficient mutants selected in *S. cerevisiae* were also defective in CV-1 cells, there were quantitative differences in the impact of these mutations. All of the mutants tested in mammalian cells retained a modest ability to induce the reporter in the presence of hormone. In addition, superactivator mutants in *S. cerevisiae* were transactivation deficient in mammalian cells. Two major factors may account for the differences in behavior of TR mutants in yeast and mammalian cells. First, because the yeast transcriptional machinery has not evolved to respond to TR, it is only reasonable that the yeast does not reflect every feature of TR in the exact manner with which mammalian cells respond. Consequently, although our screen for mutant TR in *S. cerevisiae* will detect certain residues critical for function, others not conserved in their action will be undetected. Second, and more importantly, the nuclear milieus of the two cells are different. *S. cerevisiae* does not possess RXR. As will be described below, the absence of RXR in *S. cerevisiae* might be responsible for many of the differences we observe in the behavior of mutant TR in the two cells.

The use of the yeast genetic screen allowed us to rapidly recover amino acid substitution mutations that were defective in either hormone binding or transcriptional activation. In this study we focused on a group of defective TR mutants that retained $T_3$ binding. Given the high concentration of hormone used in the screening procedure, a receptor with a reduced affinity for $T_3$ and a functional activation domain would still be expected to activate the $TRE_{pal}$ reporter. This, in fact, was borne out by the properties of the superactivator mutant S15, which bound $T_3$ only weakly but activated in a hormone-dependent fashion. Privalsky et al. (49) have also shown that v-ErbA, which has 7% of wild-type hormone-binding activity, displays hormone-dependent activity in *S. cerevisiae*. Thus, the mutants selected in our screen likely do not result from changes in hormone-binding affinity but rather represent defects in some other step in the process of transcriptional activation.

**Mutations in the D (hinge) region.** The major group of mutations leading to a defect in hormone-dependent transcriptional activation localize to a 37-amino-acid region (amino acids 179 to 215) located immediately C terminal to the DNA-binding domain. These mutations occur within the D or hinge region, a segment which is poorly conserved in the family of steroid/thyroid hormone receptors. Two distinct phenotypes in mutants with mutations in the D region can be distinguished. The first group result in a loss of binding of mutant receptors to the $TRE_{pal}$ sequence. Mutation of four different amino acids (K-179, N-190, R-194, and K-206) resulted in a loss of DNA binding (Fig. 7). The designation of K179R as the inactivating mutation for DNA binding in the double mutant 58a is based on the finding that the second mutation in this mutant, R184G, is also found in mutant 91, which binds to the $TRE_{pal}$ sequence. Likewise, the L199P mutation of the double mutant 129a is also found in mutant 44a, which binds to DNA, suggesting that R194G is the mutation leading to defective $TRE_{pal}$ binding. The finding that mutations in the D region would alter DNA binding by TR was not anticipated at the outset of this work, since this function had been attributed entirely to residues of the zinc finger-containing C domain. However, recent work suggests that the D region may play a role in this process in the thyroid/retinoid receptor subfamily.

For this group of mutations, several possibilities could explain the defect in DNA binding. The mutated TR residues could make direct DNA contacts with bases of the $TRE_{pal}$ oligonucleotide. This possibility is suggested by work on the NGFI-B orphan receptor. Wilson et al. (69) found that amino acids of NGFI-B corresponding to TRβ residues 182 to 190, which were designated the A box, influence DNA binding. This orphan receptor, unlike most members of the family, binds as a monomer to an extended octamer sequence. Residues of the A box were shown to contact the upstream bases of this response element (68, 69). One of the DNA-binding-defective mutations (at N-190) occurs within the A box, whereas the other three occur at positions surrounding it. Recently, two groups have shown that TRα has the capability of binding to DNA as a monomer (29, 58). In doing so, the DNA-binding site preference is extended in the 5′ direction by two bases from the AGGTCA core motif, suggesting that residues in the A box may contact these upstream bases. Zechel et al. (72) have confirmed that binding of TRα as a monomer required residues of the A box and positions immediately upstream of the A box in the D region. It should be pointed out, however, that in our studies, the TRE sequence used for mutant screening did not contain the optimized monomer-binding site. Fur-

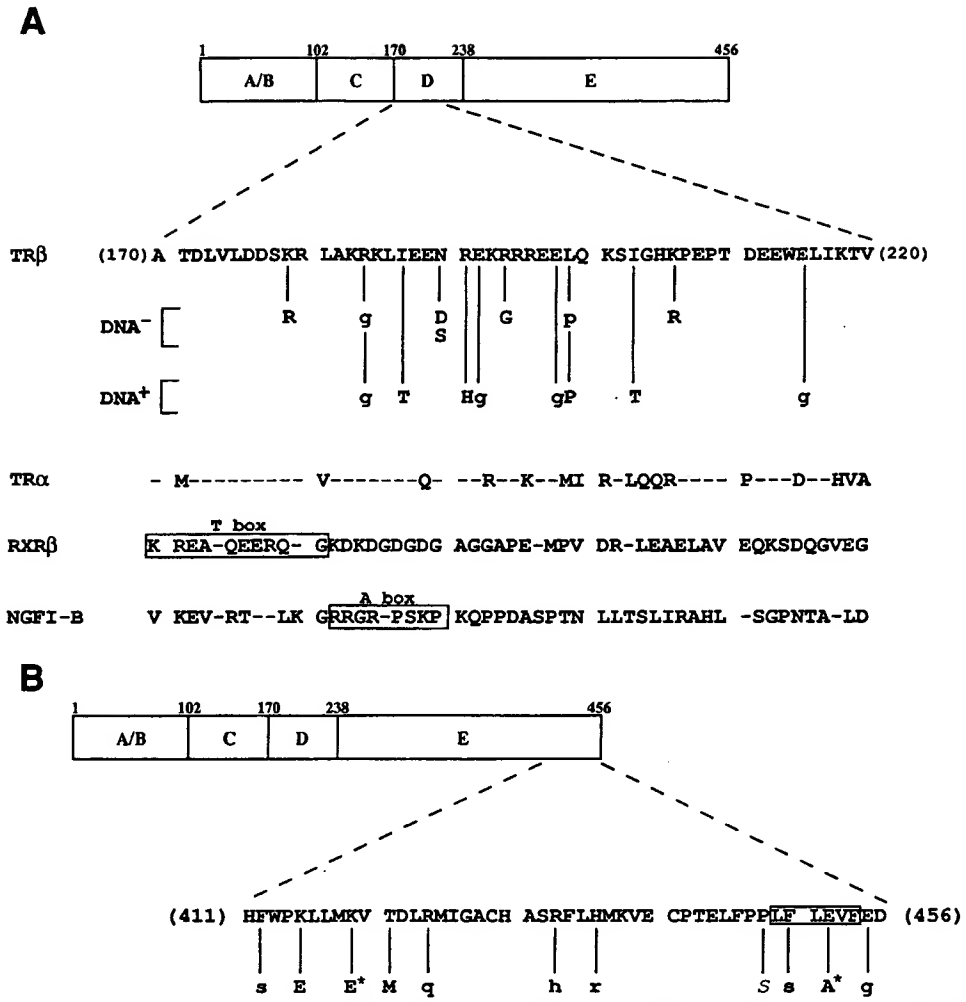



FIG. 7. (A) Sequence of the D region of TRβ and mutations in the D region leading to inactivation of receptor. Mutations are divided into those that disrupt binding to the $TRE_{pal}$ ($DNA^-$) and those that bind $TRE_{pal}$ comparably to wild-type receptor ($DNA^+$). Uppercase letters indicate mutations that can unambiguously be defined as causing the phenotype, while lowercase letters indicate mutations that are present as double mutations, for which assignment of the inactivating mutation cannot be unambiguously determined. Shown for comparison are the amino acid sequences of rat TRα, rat RXRβ, and rat NGFI-B over the homologous region. Dashes indicate residues that are identical to those in TRβ. Residues designated the T box of RXRβ and the A box of NGFI-B are indicated (69). (B) Sequence of the C-terminal end of TRβ and mutations in this region altering transcriptional activation properties of receptor. Upper- and lowercase letters are as described for panel A. Asterisks indicate residues altered in more than one mutant. Italics indicate an inactivating mutation. The box indicates a conserved region, ΦΦXEΦΦ (where Φ indicates a hydrophobic residue), in the steroid/thyroid family of receptors, as pointed out previously (7, 16).

thermore, the binding complex detected in the EMSA with the $TRE_{pal}$ probe migrates as expected for a homodimer. The question of whether the role of the A box in contacting DNA extends to homodimer binding of TR is unanswered.

A second explanation for defective DNA binding is that mutated residues could be critical for dimerization interactions in the region adjacent to the zinc fingers. This possibility was suggested by work on RXR. Residues of the hinge region were found to be critical in the binding of RXR to an RXR-specific response element (35, 69) or binding of RXR as a heterodimer with RAR to a DR+2 response element. These residues, which correspond to TRβ residues 170 to 181, were designated the T box and were suggested to play a role in dimerization. Again, we find that one of the TR mutations resulting in defective DNA binding (at K-179) is within the region corresponding to the T box.

A third possibility is that these TR mutations affect the ability of the receptor to undergo a conformational change to bring the E domains of adjacent receptors into the proper configuration for dimerization, as suggested by Kurokawa et al. (31). These workers suggested that the D region confers rotational flexibility between the C and D/E domains of TR. This rotational flexibility was postulated to account for the ability of TR to activate from TREs containing binding motifs arranged as direct repeats, palindromes, or inverted palindromes. These investigators did not test the functional activity of mutations in this region in transfection assays. However, they found that mutations at residues corresponding to the A box (181 to 190) blocked the ability of TRβ to bind to either direct repeat or inverted palindromic response elements as homodimers. At present, it is not possible to distinguish between these explanations.

In addition to the mutations in the D region that affect DNA binding, some mutations in this region also confer binding to the $TRE_{pal}$ comparable to that of the wild-type receptor yet are incapable of transactivation. Mutations in this group span the region from R-184 to E-215 and thus overlap with mutations affecting DNA binding (Fig. 7). These mutants also bind $T_3$ equivalently to the normal receptor. Thus, because both DNA binding and hormone binding are near normal, the D region must play an independent role in transcriptional activation. This conclusion is supported by the properties of GAL4 fusion proteins containing mutants in this class. Both mutants 9 and 68 diminish the activation observed with the wild-type GAL4-TR fusion protein. Since the GAL4 domain provides both DNA-binding and dimerization functions, the defects in these mutants must result from a distinct process in transcriptional activation. It is interesting that the mutant 25a, which was defective in DNA binding, also showed reduced activity as a GAL4 fusion. This result implies that this mutation affects not only DNA binding but transcriptional activation as well. Again, this result emphasizes the highly overlapping nature of the functions imparted by the D domain.

These D region mutants could affect a transactivation domain of the receptor directly involved in contacting a downstream component of the transcriptional machinery. In the glucocorticoid receptor, a hormone-independent transactivation domain, τ2, has been localized to the D region (21). However, similarly located transactivation domains have not been found in other members of the steroid receptor family. Alternatively, this region of the receptor could be important for a conformational change necessary to position the transactivation domain of TR for proper interaction with the transcriptional machinery. Several lines of evidence support the concept that hormone binding results in an alteration in receptor conformation. Toney et al. (62) have shown an altered pattern of circular dichroism associated with hormone binding. Leng et al. (37) used partial proteolytic cleavage to demonstrate that hormone binding induced a change in TR to a more protease-resistant conformation. We currently favor the possibility that these mutations in the D region interfere with this conformational change, since it may explain the overlapping nature of mutations in this region that affect DNA binding and transactivation. Both of these processes may require conformational flexibility of the D region. For DNA binding, this flexibility is necessary to properly orient the E region dimerization domains to stabilize binding, and for transactivation to allow exposure of the critical residues involved in interaction with the basal transcriptional machinery.

**Mutations dispersed in the TRβ E region.** Several of the mutants that were defective in hormone-dependent transcriptional activation contained mutations broadly dispersed in the middle of the E region. All of these mutants bound the $TRE_{pal}$ oligonucleotide comparably to the normal receptor. Several of these showed some impairment in their ability to bind $T_3$. However, we suggest that the major defect in these mutants is in their ability to transactivate.

Several mutations in this group emphasize the critical nature and overlapping role of amino acid residues between positions 280 and 310. Mutants 152 and 157 contain mutations that are close together and near a 20-amino-acid region (amino acids 281 to 300) suggested by Koenig and colleagues to be important for heterodimerization with RXR (45, 46, 51). Mutant 157 contains two changes, A226V and E306G. The latter of these is directly adjacent to this heterodimerization domain. EMSA with mutant 157 indicates that it binds with high affinity to the $TRE_{pal}$ sequence but only weakly as a heterodimer to this sequence (data not shown). This observation suggests that the heterodimerization domain may extend to at least residue 306. The properties of this mutant parallel those described for an R311H mutation in human TRβ that gives rise to a selective pituitary resistance to thyroid hormone (20). On the other hand, mutant 152 contains an L300P change and yet binds comparably to the wild-type receptor as a homodimer or RXR heterodimer (data not shown). Two of the superactivators, S19 and S20, also have mutations that are localized in this area of the receptor around residue 300. For S20, the addition of hormone actually reduces its activity. Perhaps in *S. cerevisiae* the protein is folded to expose the transactivation domain of the mutant receptor. In addition to these mutations, mutations affecting hormone binding have also been found to map to residues 304 and 308 (64). In the syndrome of GRTH, one of two hot spots for mutations is between residues 305 and 342 (50). These mutations also generally reduce hormone binding of TR. Thus, this complex region is clearly critical to multiple overlapping functions of the receptor.

Lee and Mahdavi (36) demonstrated that mutations in residues 188 to 190 of TRα1 (corresponding to residues 237 to 239 of TRβ) led to defects in transactivation but did not affect nuclear localization, $T_3$ binding, or DNA binding. Mutant 113 contains two changes (Q236P and F240S) flanking this segment and thus may alter the conformation of TR in this region. In addition, mutant 26 (L261P) is located relatively close to these residues and may influence its activity. The role of this segment in transactivation is unknown.

**Mutations in the C-terminal part of the E domain.** Several mutants that affect transcriptional activity contain changes in the last 45 amino acids of TRβ. In particular, the class of superactivator mutants includes those with mutations that are concentrated between residues F-412 and E-455 of TRβ (Fig. 7B). The most striking change occurs in mutants containing a K419E mutation, which display a 10-fold-enhanced ability to activate in the presence of hormone. Other mutants of interest in this class are S7 and S9, which contain changes of E455G and E452A, respectively. These two glutamate residues, especially E-452, are highly conserved among the steroid/thyroid receptor family, as noted by Danielian et al. (16). That mutations in this region give rise to a superactivation phenotype in *S. cerevisiae* is underscored by the fact that only a single mutant of the $T_3$-binding-competent, transactivation-deficient class with a mutation in this region was found. Mutant 5 (P448S) bound $T_3$ about 40% as well as wild-type TR yet could not activate the $TRE_{pal}$ reporter in *S. cerevisiae* either on its homologous DNA-binding domain or as a GAL4 fusion. Thus, this mutant would appear to have a minor impairment in $T_3$ binding but a major defect in its ability to transactivate. This proline is mutated to three different amino acids in separate alleles of GRTH (44).

The preponderance of superactivator mutations in a region where we might have expected to find inactivating mutations was initially puzzling. However, in retrospect, this phenotype is consistent with expected properties of mutations in the transactivation domain of a mammalian transcription factor selected in a yeast transcriptional system. We postulate that although the TR transactivation domain has been conserved in evolution to a degree sufficient to allow it to interact with the yeast basal transcriptional machinery, this interaction would not in fact be expected to be optimal. The consequence would be that TR serves as a relatively weak activator in the context of *S. cerevisiae*. Mutations that give rise to the superactivating phenotype would presumably create a stronger interaction between the transactivation domain of TR and yeast transcriptional components. This hypothesis is supported by the observation that virtually all of the superactivator mutants retain

hormone dependence for transcriptional activation. If this is true, then we might expect these same mutations to weaken the interaction with the mammalian transcriptional machinery for which TR has been evolutionarily optimized. Indeed, this is exactly the result we obtained; the mutant TR that superactivated in *S. cerevisiae* were transcriptionally crippled in mammalian cells. Thus, these data support a critical role of residues in the C-terminal region in the transactivation process.

The importance of this C-terminal region in transactivation has been supported by several studies that have appeared during the course of this work. Danielian et al. (16) noted that several C-terminal residues of the E domain were conserved across many family members. Mutation of two of these conserved positions, E-546 and D-549, of the mouse estrogen receptor seriously blunted its ligand-dependent transcriptional activation. The hormone-binding and DNA-binding capacities of this mutant were unaffected. These mutations were also introduced into the glucocorticoid receptor with similar results. On the basis of these and other data, Danielian et al. (16) propose that this region is a component of the ligand-dependent activation domain (AF-2) of estrogen and glucocorticoid receptors and probably other members of the receptor family as well. One of the dominant negative mutations of the estrogen receptor isolated by Ince et al. (26) also maps to this region. These authors suggest that the dominant negative phenotype results from AF-2 inactivation.

These results have been extended to include members of the thyroid/retinoid subfamily of receptors. Saatcioglu et al. (52) introduced simultaneous changes at glutamate residues 401 and 404 in chicken TRα1 (corresponding to E-452 and E-455 of TRβ). This mutated receptor was inactive. Other mutations in the C-terminal nine amino acids also produced receptors defective in transcriptional activation. Thus, these investigators also argued that this region must be important for transcriptional activity. However, all of the mutant TR were significantly reduced in T3 binding, as well as in transcriptional activation. Thus, this study cannot unambiguously distinguish between effects on hormone binding and on transactivation. More recently, Barettino et al. (7) extended these results for the chicken TRα by mutating several additional C-terminal residues. They found a single mutant, carrying mutation E401Q, which retained hormone-binding capacity but lost transactivation. Furthermore, they demonstrated that a fusion of the GAL4 DNA-binding domain with the 35 C-terminal residues of TRα resulted in a protein capable of activating transcription in mammalian cells. Although the level of activation was reduced significantly relative to that of a GAL4 fusion of the entire D and E domains and lacked hormone dependence, these data suggest that this region of TRα behaves as an autonomous activation domain.

As mentioned previously, mutations that affect transactivation could do so by altering intramolecular signal transduction triggered by hormone binding or by affecting intermolecular interactions between TR and the transcriptional machinery. Upon hormone binding, TR is known to undergo a conformational change (37, 62). This change presumably allows the transactivation domain of TR to productively interact with the transcriptional machinery. The characteristics of these superactivators strongly argue that the C-terminal domain of TR is directly involved in the transactivation function. If these mutations caused their phenotype by mimicking the conformational change that occurs upon hormone binding, the expected phenotype would be a hormone-independent activator. Instead, we find that virtually all of the superactivators retain a dependence on hormone binding. Thus, a hormone-dependent conformational change is necessary to expose the activation domain of the superactivator mutants. This concept is consistent with the data of Barettino et al. (7) demonstrating that this region of TRα can impart transactivation to a heterologous DNA-binding domain. However, the retention of hormone dependence for activation of the superactivator TRs more strongly supports the physiological role of this domain.

**What is the active TR species in *S. cerevisiae*?** Recent studies have suggested that TR function in mammalian cells is intimately linked with that of the RXRs. Using gel retardation assays, we have not been able to detect an RXR-like activity in yeast extracts. Similar efforts to detect a factor that enhances DNA binding of TR by using the less stringent avidin-biotin complex DNA binding assay were also unfruitful (data not shown). While we cannot completely exclude the existence of such a factor, we strongly suspect that TR functions as either a monomer or a homodimer in *S. cerevisiae*. Recently, two laboratories have shown that coexpression of RXR and TR in *S. cerevisiae* leads to enhanced TR activation, as expected from results in mammalian cells (24, 34). Consequently, we will be testing the mutant TR isolated in this study in the context of RXR in yeast cells.

An assessment of homodimer action in mammalian cells is complicated by the background of RXR found in virtually all cells. However, there is growing evidence that RXR-independent pathways of transcriptional activation do exist for members of the retinoid, thyroid, and vitamin D family of receptors. For the vitamin D receptor, RXR-independent and -dependent pathways have been defined (60). Recent work of Saatcioglu et al. (53) has shown that TRα can function as a hormone-independent activator from a unique Rous sarcoma virus DNA element. This action is not dependent on RXR. Kurokawa et al. (31) presented evidence that TR were able to form stable homodimers in solution. The RXR-TR interaction was approximately threefold higher than the TR-TR interaction in this assay. Katz and Koenig (29) used a PCR-based binding site selection strategy to identify an 8-bp element (5′-TAAGGTCA) that binds strongly to TRα monomers. This element also confers TR-dependent activation to a linked reporter gene in cotransfection assays. Recent work of Schrader et al. (59) confirms these findings. All of these observations suggest that in addition to heterodimers, homodimers and monomers of TR expand the repertoire of gene regulation by the thyroid hormones.

In conclusion, we have identified specific amino acid residues in TR which play a critical role in the process of hormone-dependent activation in yeast and mammalian cells. In particular, two regions of TRβ contained multiple mutations: the D region and a carboxy-terminal segment of the E region. Our data suggest that mutations in these residues affect the process of transcriptional activation, independent from hormone binding. These mutants should prove useful in further delineating the pathway of action of TR in mammalian cells.

## ACKNOWLEDGMENTS

We thank Martin Privalsky (University of California, Davis) for providing the original yeast expression and reporter vectors and Mingjer Tsai (Baylor College of Medicine) for pLGSD5. We thank Anupam Kharbanda for excellent technical assistance. Special thanks are also extended to Dennis Livingston, Paul Siliciano, Ken Dornfeld, and Mike Kaytor for advice on the yeast system.

This work was supported by a grant from the National Institutes of Health (DK39997). R.U. was supported in part by National Institutes of Health training grant 5T32-GM07323.

## REFERENCES

1. **Ahn, B.-Y., and D. M. Livingston.** 1986. Mitotic gene conversion-lengths, coconversion patterns, and the incidence of reciprocal recombination in a *Saccharomyces cerevisiae* plasmid system. Mol. Cell. Biol. 6:3685–3693.

2. **Andersson, S., D. L. Davis, H. Dahlback, H. Jornvall, and D. W. Russell.** 1989. Cloning, structure, and expression of the mitochondrial cytochrome P-450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme. J. Biol. Chem. **264**:8222–8229.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley and Sons, New York.
4. **Baniahmad, A., I. Ha, D. Reinberg, S. Tsai, M.-J. Tsai, and B. W. O'Malley.** 1993. Interaction of human thyroid hormone receptor β with transcription factor TFIIB may mediate target gene derepression and activation by thyroid hormone. Proc. Natl. Acad. Sci. USA **90**:8832–8836.
5. **Baniahmad, A., C. Steiner, A. C. Kohne, and R. Renkawitz.** 1990. Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. Cell **61**:505–514.
6. **Baniahmad, A., S. Y. Tsai, B. W. O'Malley, and M. J. Tsai.** 1992. Kindred S thyroid hormone receptor is an active and constitutive silencer and a repressor for thyroid hormone and retinoic acid responses. Proc. Natl. Acad. Sci. USA **89**:10633–10637.
7. **Barettino, D., M. D. M. Vivanco Ruiz, and H. G. Stunnenberg.** 1994. Characterization of the ligand-dependent transactivation domain of the thyroid hormone receptor. EMBO J. **13**:3039–3049.
8. **Becker, D. M., and L. Guarente.** 1991. High-efficiency transformation of yeast by electroporation. Methods Enzymol. **194**:182–187.
9. **Bodenner, D. L., M. A. Mroczynski, B. D. Weintraub, S. Radovick, and F. E. Wondisford.** 1991. A detailed functional and structural analysis of a major thyroid hormone inhibitory element in the human thyrotropin β-subunit gene. J. Biol. Chem. **266**:21666–21673.
10. **Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink.** 1985. Ty elements transpose through an RNA intermediate. Cell **40**:491–500.
11. **Brent, G. A., M. K. Dunn, J. W. Harney, T. Gulick, P. R. Larsen, and D. D. Moore.** 1989. Thyroid hormone aporeceptor represses T3-inducible promoters and blocks activity of the retinoic acid receptor. New Biol. **1**:329–336.
12. **Cama, A., B. Marcus-Samuels, and S. I. Taylor.** 1988. Immunological abnormalities in insulin receptors on cultured EBV-transformed lymphocytes from insulin-resistant patient with leprechaunism. Diabetes **37**:982–988.
13. **Casanova, J., E. Helmer, S. Selmi-Ruby, J.-S. Qi, M. Au-Fliegner, V. Desai-Yajnik, N. Koudinova, F. Yarm, B. M. Raaka, and H. H. Samuels.** 1994. Functional evidence for ligand-dependent dissociation of thyroid hormone and retinoic acid receptors from an inhibitory cellular factor. Mol. Cell. Biol. **14**:5756–5765.
14. **Damm, K., and R. M. Evans.** 1993. Identification of a domain required for oncogenic activity and transcriptional supression by v-erbA and thyroid hormone receptor α. Proc. Natl. Acad. Sci. USA **90**:10668–10672.
15. **Damm, K., C. C. Thompson, and R. M. Evans.** 1989. Protein encoded by v-erbA functions as a thyroid hormone receptor antagonist. Nature (London) **339**:593–597.
16. **Danielian, P. S., R. White, J. A. Lees, and M. G. Parker.** 1992. Identification of a conserved region required for hormone dependent transcriptional activation by steroid receptors. EMBO J. **11**:1025–1033.
17. **Evans, R. M.** 1988. The steroid and thyroid hormone receptor superfamily. Science **240**:889–895.
18. **Fondell, J. D., A. L. Roy, and R. G. Roeder.** 1993. Unliganded thyroid hormone receptor inhibits formation of a functional preinitiation complex: implications for active repression. Genes Dev. **7**:1400–1410.
19. **Garabedian, M. J., and K. R. Yamamoto.** 1992. Genetic dissection of the signaling domain of a mammalian steroid receptor in yeast. Mol. Biol. Cell **3**:1245–1257.
20. **Geffner, M. E., F. Su, N. S. Ross, J. M. Hershman, C. Van Dop, J. B. Menke, E. Hao, R. K. Stanzak, T. Eaton, H. H. Samuels, and S. J. Usala.** 1993. An arginine to histidine mutation in codon 311 of the c-erbA gene results in a mutant thyroid hormone receptor that does not mediate a dominant negative phenotype. J. Clin. Invest. **91**:538–546.
21. **Giguere, V., S. M. Hollenberg, M. G. Rosenfeld, and R. M. Evans.** 1986. Functional domains of the human glucocorticoid receptor. Cell **46**:645–652.
22. **Glass, C. K., J. M. Holloway, O. V. Devary, and M. G. Rosenfeld.** 1988. The thyroid hormone receptor binds with opposite transcriptional effects to a common sequence motif in thyroid hormone and estrogen response elements. Cell **54**:313–323.
23. **Guarente, L., R. R. Yocum, and P. Gifford.** 1982. A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. Proc. Natl. Acad. Sci. USA **79**:7410–7414.
24. **Hall, B. L., Z. Smit-McBride, and M. L. Privalsky.** 1993. Reconstitution of retinoid X receptor function and combinatorial regulation of other nuclear hormone receptors in the yeast *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **90**:6929–6933.
25. **Himmelfarb, H. J., J. Pearlberg, D. H. Last, and M. Ptashne.** 1990. GAL11P: a yeast mutation that potentiates the effect of weak GAL4-derived activators. Cell **63**:1299–1309.
26. **Ince, B. A., Y. Zhuang, C. K. Wrenn, D. J. Shapiro, and B. S. Katzenellenbogen.** 1993. Powerful dominant negative mutants of the human estrogen receptor. J. Biol. Chem. **268**:14026–14032.
27. **Inoue, A., J. Yamakawa, M. Yukioka, and S. Morisawa.** 1983. Filter-binding assay procedure for thyroid hormone receptors. Anal. Biochem. **134**:176–183.
28. **Jacoby, D. B., N. D. Zilz, and H. C. Towle.** 1989. Sequences within the 5′-flanking region of the $S_{14}$ gene confer responsiveness to glucose in primary hepatocytes. J. Biol. Chem. **264**:17623–17626.
29. **Katz, R., and R. Koenig.** 1993. Nonbiased identification of DNA sequences that bind thyroid hormone receptor alpha1 with high affinity. J. Biol. Chem. **268**:19392–19397.
30. **Krust, A., S. Green, P. Argos, V. Kumar, P. Walter, J. M. Bornert, and P. Chambon.** 1986. The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors. EMBO J. **5**:891–897.
31. **Kurokawa, R., V. C. Yu, A. Naar, S. Kyakumoto, Z. Han, S. Silverman, M. G. Rosenfeld, and C. K. Glass.** 1993. Differential orientations of the DNA-binding domain and carboxy-terminal dimerization interface regulate binding site selection by nuclear receptor heterodimers. Genes Dev. **7**:1423–1435.
32. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
33. **Lazar, M. A.** 1993. Thyroid hormone receptors: multiple forms, multiple possibilities. Endocrine Rev. **14**:184–193.
34. **Lee, J. W., D. D. Moore, and R. A. Heyman.** 1994. A chimeric thyroid hormone receptor constitutively bound to DNA requires retinoid X receptor for hormone-dependent transcriptional activation in yeast. Mol. Endocrinol. **8**:1245–1252.
35. **Lee, M. S., S. A. Kliewer, J. Provencal, P. E. Wright, and R. M. Evans.** 1993. Structure of the retinoid X receptor α DNA binding domain: a helix required for homodimeric DNA binding. Science **260**:1117–1121.
36. **Lee, Y., and V. Mahdavi.** 1993. The D domain of the thyroid hormone receptor α1 specifies positive and negative transcriptional regulation functions. J. Biol. Chem. **268**:2021–2028.
37. **Leng, X., S. Y. Tsai, B. W. O'Malley, and M.-J. Tsai.** 1993. Ligand-dependent conformational changes in thyroid hormone and retinoic acid receptors are potentially enhanced by heterodimerization with retinoic X receptor. J. Steroid Biochem. Mol. Biol. **46**:643–661.
38. **Leung, D. W., E. Chen, and D. V. Goeddel.** 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique **1**:11–15.
39. **Liu, H.-C., and H. C. Towle.** 1994. Functional synergism between multiple thyroid hormone response elements regulates hepatic expression of the rat $S_{14}$ gene. Mol. Endocrinol. **8**:1021–1037.
40. **Mader, S., V. Kumar, H. de Verneuil, and P. Chambon.** 1989. Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element. Nature (London) **338**:271–274.
41. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371.
42. **Mixson, A. J., R. Parrilla, S. C. Ransom, E. A. Wiggs, J. H. McClaskey, P. Hauser, and B. D. Weintraub.** 1992. Correlations of language abnormalities with localization of mutations in the β-thyroid hormone receptor in 13 kindreds with generalized resistance to thyroid hormone: identification of four new mutations. J. Clin. Endocrinol. Metab. **75**:1039–1045.
43. **Murray, M. B., N. D. Zilz, N. L. McCreary, M. J. MacDonald, and H. C. Towle.** 1988. Isolation and characterization of rat cDNA clones for two distinct thyroid hormone receptors. J. Biol. Chem. **263**:12770–12777.
44. **Nagaya, T., and J. L. Jameson.** 1993. Thyroid hormone receptor dimerization is required for dominant negative inhibition by mutations that cause thyroid hormone resistance. J. Biol. Chem. **268**:15766–15771.
45. **O'Donnell, A. L., and R. J. Koenig.** 1990. Mutational analysis identifies a new functional domain of the thyroid hormone receptor. Mol. Endocrinol. **4**:715–720.
46. **O'Donnell, A. L., E. D. Rosen, D. S. Darling, and R. J. Koenig.** 1991. Thyroid hormone receptor mutations that interfere with transcriptional activation also interfere with receptor interaction with a nuclear protein. Mol. Endocrinol. **5**:94–99.
47. **Ohashi, H., Y. F. Yang, and P. G. Walfish.** 1991. Rat liver c-erbA beta 1 thyroid hormone receptor is a constitutive activator in yeast (Saccharomyces cerevisiae): essential role of domains C, E, and F in hormone-independent transcription. Biochem. Biophys. Res. Commun. **178**:1167–1175.
48. **Olesen, J., S. Hahn, and L. Guarente.** 1987. Yeast HAP2 and HAP3 activators both bind to the CYC1 upstream activating site, UAS2, in an interdependent manner. Cell **51**:953–961.
49. **Privalsky, M. L., M. Sharif, and K. R. Yamamoto.** 1990. The viral erbA oncogene protein, a constitutive repressor in animal cells, is a hormone-regulated activator in yeast. Cell **63**:1277–1286.
50. **Refetoff, S.** 1993. Resistance to thyroid hormone. Clin. Lab. Med. **13**:563–581.
51. **Rosen, E. D., E. G. Beninghof, and R. J. Koenig.** 1993. Dimerization interfaces of thyroid hormone, retinoic acid, vitamin D, and retinoid X receptors. J. Biol. Chem. **268**:11534–11541.
52. **Saatcioglu, F., P. Bartunek, T. Deng, M. Zenke, and M. Karin.** 1993. A conserved C-terminal sequence that is deleted in v-ErbA is essential for the biological activities of c-ErbA (the thyroid hormone receptor). Mol. Cell. Biol. **13**:3675–3685.

53. **Saatcioglu, F., T. Deng, and M. Karin.** 1993. A novel cis element mediating ligand-independent activation of c-erbA: implications for hormonal regulation. Cell **75**:1095–1105.
54. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
55. **Samuels, H. H., F. Stanley, and J. Casanova.** 1979. Depletion of L-3,5,3′-triiodothyronine and L-thyroxine in euthyroid calf serum for use in cell culture studies of the action of thyroid hormone. Endocrinology **105**:80–85.
56. **Schena, M., L. P. Freedman, and K. R. Yamamoto.** 1989. Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. Genes Dev. **3**:1590–1601.
57. **Schena, M., K. Picard, and K. R. Yamamoto.** 1991. Vectors for constitutive and inducible gene expression in yeast. Methods Enzymol. **194**:389–397.
58. **Schrader, M., M. Becker-Andre, and C. Carlberg.** 1994. Thyroid hormone receptor functions as monomeric ligand-induced transcription factor on octameric half-sites. J. Biol. Chem. **269**:6444–6449.
59. **Schrader, M., K. M. Muller, and C. Carlberg.** 1994. Specificity and flexibility of vitamin D signaling. J. Biol. Chem. **269**:5501–5504.
60. **Schrader, M., A. Wyss, L. J. Sturzenbecker, J. F. Grippo, P. LeMotte, and C. Carlberg.** 1993. RXR-dependent and RXR-independent transactivation by retinoic acid receptors. Nucleic Acids. Res. **5**:1231–1237.
61. **Thompson, C. C., and R. M. Evans.** 1989. Trans-activation by thyroid hormone receptors: functional parallels with steroid hormone receptors. Proc. Natl. Acad. Sci. USA **86**:3494–3498.
62. **Toney, J. H., L. Wu, A. E. Summerfield, G. Sanyal, B. M. Forman, J. Zhu, and H. H. Samuels.** 1993. Conformation changes in chicken thyroid hormone receptor α1 induced by binding to ligand or to DNA. Biochemistry **32**:2–6.
63. **Umesono, K., and R. M. Evans.** 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell **57**:1139–1146.
64. **Uppaluri, R., A. Kharbanda, and H. C. Towle.** 1994. Unpublished data.
65. **Vegeto, E., G. F. Allan, W. T. Schrader, M. J. Tsai, D. P. McDonnell, and B. W. O'Malley.** 1992. The mechanism of RU486 antagonism is dependent on the conformation of the carboxy-terminal tail of the human progesterone receptor. Cell **69**:703–713.
66. **Webster, N. J. G., S. Green, J. R. Jin, and P. Chambon.** 1988. The hormone-binding domains of the estrogen and glucocorticoid receptors contain an inducible transcription activation function. Cell **54**:199–207.
67. **Weiss, R. E., M. Weinberg, and S. Refetoff.** 1993. Identical mutations in unrelated families with generalized resistance to thyroid hormone occur in cytosine-guanine-rich areas of the thyroid hormone receptor beta gene. J. Clin. Invest. **91**:2408–2415.
68. **Wilson, T. E., T. J. Fahrner, and J. Milbrandt.** 1993. The orphan receptors NGFI-B and steroidogenic factor 1 establish monomer binding as a third paradigm of nuclear receptor-DNA interaction. Mol. Cell. Biol. **13**:5794–5804.
69. **Wilson, T. E., R. E. Paulsen, K. A. Padgett, and J. Milbrandt.** 1992. Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. Science **256**:107–110.
70. **Wrenn, C. K., and B. S. Katzenellenbogen.** 1993. Structure-function analysis of the hormone binding domain of the human estrogen receptor by region-specific mutagenesis and phenotypic screening in yeast. J. Biol. Chem. **268**:24089–24098.
71. **Yen, P. M., A. Sugawara, and W. W. Chin.** 1992. Triiodothyronine differentially affects T3-receptor/retinoic acid receptor and T3-receptor/retinoid X receptor heterodimer binding to DNA. J. Biol. Chem. **267**:23248–23252.
72. **Zechel, C., X.-Q. Shen, P. Chambon, and H. Gronemeyer.** 1994. Dimerization interfaces formed between the DNA binding domains determine the cooperative binding of RXR/RAR and RXR/TR heterodimers to DR5 and DR4 elements. EMBO J. **13**:1414–1424.
73. **Zilz, N. D., M. B. Murray, and H. C. Towle.** 1990. Identification of multiple thyroid hormone response elements in the far upstream region of the $S_{14}$ gene. J. Biol. Chem. **265**:8136–8143.

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 271, No. 9, Issue of March 1, pp. 4872–4878, 1996
Printed in U.S.A.

# Human Immunodeficiency Virus Reverse Transcriptase

## FUNCTIONAL MUTANTS OBTAINED BY RANDOM MUTAGENESIS COUPLED WITH GENETIC SELECTION IN *ESCHERICHIA COLI**

(Received for publication, November 10, 1995, and in revised form, December 20, 1995)

**Baek Kim, Tanya R. Hathaway, and Lawrence A. Loeb‡**

*From The Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology, Box 357705, University of Washington, Seattle, Washington 98195-7705*

**We describe catalytically active mutants of HIV RT (human immunodeficiency virus reverse transcriptase) generated by random sequence mutagenesis and selected in *Escherichia coli* for ability to complement the temperature-sensitive phenotype of a DNA polymerase I (Pol I$^{ts}$) mutant. We targeted amino acids Asp-67 through Arg-78 in HIV RT, which form part of the β3-β4 flexible loop and harbor many of the currently known mutations that confer resistance to nucleoside analogs. DNA sequencing of 109 selected mutants that complement the Pol I$^{ts}$ phenotype revealed substitutions at all 12 residues targeted, indicating that none of the wild-type amino acids is essential. However, single mutations were not observed at Trp-71, Arg-72, and Arg-78, consistent with evolutionary conservation of these residues among viral RTs and lack of variation at these positions among isolates from patients. The mutations we recovered included most of those associated with drug resistance as well as previously unidentified mutations. Purification and assay of 14 mutant proteins revealed correlation between their DNA-dependent DNA polymerize activity *in vitro* and ability to complement the Pol I$^{ts}$ phenotype. Activity of several mutants was resistant to 3′-azidothymidine triphosphate. We conclude that random sequence mutagenesis coupled with positive genetic selection in *E. coli* yields large numbers of functional HIV RT mutants. Among these are less active variants which are unlikely to be isolated from HIV-infected individuals and which will be informative of the roles of individual amino acids in the catalytic functions of the enzyme.**

Infection by human immunodeficiency virus 1 (HIV-1)[1] is the pathogenic precursor to clinical development of acquired immunodeficiency syndrome (AIDS). HIV-1 contains a reverse transcriptase (HIV RT) that catalyzes synthesis of both a single- and double-stranded DNA copy of the viral genome, and is required for viral replication. As such, HIV RT provides a central target for chemotherapy of AIDS. Current chemotherapeutic agents that target HIV RT, such as AZT, ddC, and ddI, are incorporated into viral DNA but cannot be further elongated, thus terminating DNA synthesis (Mitsuya and Broder, 1986). Unfortunately, the exceptionally high mutation rate of the virus drives rapid emergence of resistant strains, rendering these nucleoside analogs ineffective (Boucher *et al.*, 1992; Larder and Kemp, 1989; Mitsuya and Broder, 1986). Single and multiple base substitutions within the HIV RT gene are responsible for emergence of these drug-resistant mutants that escape chemotherapy. In order to develop drugs that more effectively target HIV RT, we require a clearer understanding of the potential of different sites within the enzyme to produce drug-resistant mutations.

Many of the mutations that render HIV resistant to nucleoside analogs are located in the β3-β4 loop, β9-β10 turn, and β11a/b regions of HIV RT (Tantillo *et al.*, 1994). The β3-β4 region in the fingers domain is believed to function as a template grip that interacts with single-stranded (ss) DNA and RNA templates (Boyer *et al.*, 1994; Tantillo *et al.*, 1994). Mutations within the β3-β4 loop that confer drug resistance include K65R (resistant to ddC, ddI, and 3TC (Gu *et al.*, 1994)), D67N (resistant to AZT (Larder and Kemp, 1989), Larder and Kemp, 1989)), T69D (resistant to ddC (Fitzgibbon *et al.*, 1992)), K70R (resistant to AZT (Larder and Kemp, 1989)), and L74V (resistant to ddI (St. Clair *et al.*, 1991)). Recently, two additional mutations, V75I and F77L, have been isolated from patients treated with dideoxynucleosides (Shirasaka *et al.*, 1995). The diverse locations of mutations in HIV RT that render the virus resistant to the same drug has prompted the suggestion that many of these mutations act indirectly, at a distance. It is postulated that such mutations affect conformation near the catalytic site, modifying interactions between the fingers domain and the template DNA or RNA (Boyer *et al.*, 1994; Tantillo *et al.*, 1994) to enhance discrimination between natural substrates and inhibitors. Other drug-resistant mutations, not in β3-β4 loop region, such as M184V (Gu *et al.*, 1992), that are located close to the catalytic site are proposed to directly affect dNTP binding (Tantillo *et al.*, 1994).

We recently demonstrated that HIV RT can substitute for DNA polymerase I in *Escherichia coli* (Kim and Loeb, 1995a). In our genetic complementation system, expression of HIV RT allows mutant *E. coli* cells harboring a temperature-sensitive DNA polymerase I (Pol I$^{ts}$) to grow at non-permissive temperature. Pol I$^{ts}$ cells expressing an inactive mutant (D186N) fail to grow at nonpermissive temperature, demonstrating the requirement for catalytically functional HIV RT. Based on the observation that complementation by HIV RT is specifically inhibited by nucleoside analogs such as AZT, we proposed that the system can be used to screen potential anti-HIV RT drugs in bacteria (Kim and Loeb, 1995b).

In the present study, we used the bacterial complementation

* This work was supported by National Institutes of Health Grant R35-CA-39903 (to L. A. L.) and by National Institutes of Health Molecular Training Grant 5T32-CA-09437-13 (to B. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 206-543-0557; Fax: 206-543-3967.

[1] The abbreviations used are: HIV RT, human immunodeficiency virus reverse transcriptase; AZT, 3′-azidothymidine; AZTTP, 3′-azidothymidine triphosphate; ddC, 2′,3′-dideoxycytidine; ddI, 2′,3′-dideoxyinosine; dNTP, deoxynucleoside triphosphate; IPTG, isopropyl-1-thio-β-D-galactopyranoside; $IC_{50}$, concentration yielding 50% inhibition; Pol I$^{ts}$, temperature-sensitive polymerase I.

system to select active HIV RT mutants from libraries that encode random sequence substitutions in the wild type gene. We replaced the 36 contiguous nucleotides near the β3-β4 loop that encode amino acids 67 through 78 with random sequences and selected from this library mutants that complement the Pol I$^{ts}$ mutation. This positive genetic selection allowed us to recover most of the active HIV RT mutations thus far observed in naturally occurring variants and drug-resistant mutants. Our large library of active mutants includes previously unidentified mutations and mutations in the β3-β4 loop that confer resistance to AZTTP.

## EXPERIMENTAL PROCEDURES

*Strains*—*E. coli* NM522 (Stratagene) was used for cloning and library construction. *E. coli* BL21(DE3) from Novagen was used for expression and purification of HIV RT. 2 × YT (Sambrook *et al.*, 1989) was used for growth of these strains. The Pol I$^{ts}$ strain used for genetic complementation is the *E. coli* B/r derivative SC18-12 (*rec*A718 *pol*A12 *uvr*155 *trp*E65 *lon*-11 *sul*A1) (Witkin and Roegner, 1992). Nutrient broth containing Difco nutrient broth (8 g/liter), NaCl (4 g/liter), and tetracycline (12.5 μg/ml) was used for culturing Pol I$^{ts}$ cells. Nutrient agar plates containing Difco nutrient agar (23 g/liter), NaCl (5 g/liter), and tetracycline (12.5 μg/ml) were used for assessing genetic complementation.

*Plasmids*—pHIVRT is a derivative of pHSG576 that contains the wild-type HIV RT under control of *lac*P/O (Kim and Loeb, 1995a). pBK8 is a derivative of pBS-SK(+) (Stratagene) containing the HIV RT gene generated from digestion of pHIVRT with *Hin*dIII and *Eco*RI. A *Sac*I site at position 79 of the HIV RT gene in pBK8 was created by site-directed mutagenesis (Kunkel, 1985) to generate pBK13. To avoid mutations that might be created in other regions of the HIV RT gene during site-directed mutagenesis, the 310-bp segment of HIV RT in pBK13 flanked by *Bsr*GI and *Eco*RV restriction sites was replaced with the corresponding segment of pBK8 containing the *Sac*I site to generate pBK14. The 1.3-kilobase wild-type sequence between the *Bsr*GI and *Eco*RI sites of the HIV RT gene in pHIVRT was replaced with the corresponding 1.3-kilobase fragment generated by digestion of pBK14 with *Bsr*GI and *Eco*RI, generating pBK16. Thus, pBK16 is the same as pHIVRT, but contains a *Sac*I site at position 79 of the HIV RT gene. To avoid contamination with wild-type pHIVRT during genetic selection, a "stuffer" plasmid was constructed in which the 108-bp wild-type sequence of pBK16 between *Bsr*GI and *Sac*I was replaced with a 3,515-bp DNA fragment by digestion of phage λ DNA with *Bsr*GI and *Sac*I, followed by ligation, to generate pBK18 (*step 5*, Fig. 1). Chloramphenicol (30 μg/ml) was used for selection of the pHSG576 derivatives pHIVRT, pBK16, and pBK18 (Takeshita *et al.*, 1987). Carbenicillin (50 μg/ml) was used for maintenance of the pBR322 derivatives pBK8, pBK13, and pBK14.

*Construction of HIV RT Random Library*—The HIV RT random library was constructed by annealing two single-stranded DNA oligomers. Oligomer 1 (Fig. 1, *steps 1* and *2*) was the 83-mer (5′-GTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGC-3′) corresponding to the plus strand of the HIV RT sequence from amino acid residues 35 to 62. Oligomer 2 was the 79-mer (5′-CTTATTGAGCTCTCTGAAATCTACTAATTTTCTCCATTTAGTACTGTCTTTTTTCTTTATGGCAAATACTGGAGTATTG-3′) corresponding to the minus strand of the HIV RT sequence from amino acids 56 to 82. Oligomer 2 contained mutated sequences corresponding to amino acids 67 to 78 (underlined), with 12% random nucleotides and 88% wild-type nucleotide at each of the 36 randomized positions, and thus encoded an average of 4.3 nucleotide substitutions per primer. The two primers (1 μg each) were mixed in 50 μl of 200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl (5 × sequencing buffer (U. S. Biochemical Corp)) and annealed by incubating at 80 °C for 5 min, at 55 °C for 15 min, at 37 °C for 15 min, and at room temperature for 15 min (*step 1*, Fig. 1). The partially double-stranded oligonucleotide was extended by incubation with Sequenase Version 2.0 T7 DNA polymerase (U. S. Biochemical Corp., 0.6 unit/ml) and dNTPs (1 mM) for 2 h at 37 °C in a total volume of 54 μl (*step 2*, Fig. 1). The double-stranded oligonucleotides (2 μl) were then amplified in a polymerase chain reaction using Vent DNA polymerase (New England Biolabs) and 5′ and 3′ end primers in 100 μl of 20 mM Tris-HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100 at an annealing temperature of 28 °C for 30 cycles (*step 3*, Fig. 1). The 5′ primer was a 14-mer identical to the 5′ end of oligomer 1, and the 3′ primer was a 14-mer identical to the 5′ end of oligomer 2. The amplified DNA (5 μg) was digested with *Bsr*GI and *Sac*I (New England Biolabs) (*step 4*, Fig. 1), purified, and concentrated by centrifugation with a Microcone 30 filter (Amicon), and phenol extracted. The purified oligonucleotides were then used as the random nucleotide-containing inserts for construction of the HIV RT plasmid library. The stuffer plasmid pBK18 (10 μg) was also digested with *Bsr*GI and *Sac*I. The resulting 5.0-kilobase fragment was isolated from a 0.8% agarose gel by using an Ultrafree-MC 0.45-mm filter (Millipore) and 5 μg was ligated to the 108-bp insert by using T4 DNA ligase (Boehringer Mannheim) (*step 5*, Fig. 1). The ligation mixture was directly transformed into electrocompetent NM522 cells (Stratagene) (*step 6*, Fig. 1). A plasmid library was prepared from an overnight culture of transformed cells grown in 2 × YT containing 30 μg/ml chloramphenicol (*step 7*, Fig. 1).

*Genetic Selection in E. coli*—Electroporation of a 50-μl mixture of 8 × $10^9$ Pol I$^{ts}$ cells and 2 μg of plasmid library was carried out by using a Gene Pulsar apparatus (Bio-Rad; 25 microfarads, 2.1 V, 400 ohms). Transformed cells were plated onto nutrient agar selection plates containing chloramphenicol, tetracycline, and IPTG (1 mM) (about 200 cells per plate; *step 8*, Fig. 1). Transformation efficiency, measured after incubation at permissive temperature (30 °C) for 24 h, was typically 2 × $10^7$ cells per 1 μg of DNA. For selection of active HIV RT mutants, plates were incubated at nonpermissive temperature (37 °C) for 18 h. Colonies obtained at 37 °C were regrown individually in 2 ml of nutrient broth containing chloramphenicol and tetracycline. To verify the phenotype, plasmids were purified from each of these cultures and re-transformed into Pol I$^{ts}$ cells by electroporation. About 200 re-transformed cells were plated onto each of two nutrient agar selection plates, followed by incubation of one plate at 30 °C and the other at 37 °C. After incubation for 18 h at 37 °C, the extent of complementation by each of the selected mutants was scored visually with respect to colony size (+ through +++, with +++ being equivalent to wild-type) and plating efficiency (number of colonies at 37 °C/number of colonies at 30 °C).

*DNA Sequencing*—Plasmid DNA, purified from Pol I$^{ts}$ cells, was transformed into *E. coli* NM522. Approximately 0.4 μg of plasmid DNA was prepared from NM522 cells. The nucleotide sequence encoding amino acids 37 to 86 of HIV RT and including the randomized segment was established by polymerase chain reaction-based sequencing (fmol sequencing kit, Promega). The sequencing primer (5′-TTAAACAATGGCCATTGACAG) was a 22-mer encoding amino acid residues 21 to 28 of HIV RT.

*Purification of Wild-type HIV RT*—To construct a plasmid expressing wild-type HIV RT, a 1.6-kilobase DNA fragment of pBK14 was digested with *Nde*I and *Sal*I and cloned into pET28a (Novagene) to generate pBK34. pBK34 expresses wild-type HIV RT fused at its N terminus to 20 amino acids that includes a hexahistidine sequence. The HIV RT fusion protein was purified in a one-step metal chelation chromatographic procedure by using $Ni^{2+}$ affinity resin and buffers (His Bind Resin and Buffer Kit, Novagen) according to a protocol modified from that of the supplier. Expression of HIV RT was induced by addition of 0.1 mM IPTG to 1 liter of log phase *E. coli* BL21 (DE3) grown in 2 × YT to an $OD_{600}$ of 0.5. After incubation for 2 h, cells were harvested by centrifugation and the pellets were resuspended and frozen (−70 °C) with 30 ml of 1 × binding buffer and lysozyme (200 μg/ml, Sigma). Frozen cells were thawed and lysed on ice for approximately 2 h. The lysed cells were centrifuged (27,000 × *g*) and the supernatant solution was applied to a charged 10-ml His Bind column (1 × 10 cm). The resin was prepared by successive washes with deionized water (30 ml), 1 × charge buffer (30 ml), and 1 × binding buffer (45 ml). All chromatographic steps were carried out at 4 °C at a flow rate of 20 ml/h. Following application of the crude supernatant solution, the column was washed with 1 × binding buffer (30 ml) and 1 × wash buffer (10 ml). HIV RT was eluted with 1 × elute buffer (30 ml); 90% of the recovered HIV RT was released from the resin in the first 8 ml. Fractions containing purified HIV RT were analyzed by electrophoresis in a 12% SDS-polyacrylamide gel. Fractions containing HIV RT were combined (5 ml) and dialyzed against 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 200 mM NaCl, 10% glycerol for 16 h and then against the same buffer containing 1 mM dithiothreitol for an additional 16 h. The purity of HIV RT was greater than 95%, estimated by visual inspection of Coomassie Blue-stained gels. The concentration of purified HIV RT was determined by the dye-binding procedure of Bradford (Bradford, 1976) with the dye reagent and bovine serum albumin standard from Bio-Rad.

*Purification of Mutant HIV RTs*—The 108-bp fragments containing mutations in HIV RT were inserted into pBK34 between the *Bsr*GI and *Sac*I sites as follows. The DNA fragments between the *Bsr*GI and *Sac*I sites were polymerse chain reaction-amplified with the 5′ and 3′ end primers used in Fig. 1 (*step 3*). The amplified products were digested with *Bsr*GI and *Sac*I and inserted into pBK34 digested with the same

enzymes. Expression and purification of the mutant proteins were essentially as described for wild-type HIV RT. Mutant HIV RTs were purified from 200 ml of 2 × YT cultures on a 1-ml column; the volume of all buffers was one-fifth that specified for chromatography of wild-type HIV RT. Protein concentrations were estimated by visually comparing intensities of the 66- and 51-kDa bands with those of known amounts of wild-type HIV RT in Coomassie Blue-stained 12% SDS-polyacrylamide gels.

*DNA Polymerase Activity*—DNA-dependent DNA polymerase activity was assayed with a gapped DNA template. Purified HIV RT protein (0.1 μg) was incubated in a 50-μl reaction mixture containing activated calf thymus DNA (80 μg/ml), 50 mM Tris-HCl, pH 8.0, 50 mM KCl, 8 mM $MgCl_2$, 5 mM dithiothreitol, 100 μg/ml bovine serum albumin, 5 μM each dATP, dCTP, dGTP, and dTTP, and 50 nM [*methyl*-$^3$H]dTTP (40 Ci/mmol, Moravek Biochemicals, Inc.). Incubation was at 37 °C for 20 min and incorporation was determined as described previously (Battula and Loeb, 1974). Incorporation of label from [*methyl*-$^3$H]dTTP was a linear function of amount of enzyme and incubation time.

## RESULTS

*Selection for Functional Mutants by Genetic Complementation*—To construct an HIV RT plasmid library, we substituted a segment of the wild-type HIV RT gene with an oligomer containing 12% random nucleotides and 88% wild-type nucleotide at each of the 36 positions encoding amino acids 67–78 (Fig. 1, *steps 1–4*). The partially random inserts were then ligated into pBK18, generating approximately 5 × $10^5$ plasmid-borne HIV RT mutants (Fig. 1, *step 5*). To identify active mutants, we employed a positive genetic selection system in which HIV RT complements the temperature-sensitive phenoytpe of an *E. coli* DNA polymerase I mutant (Kim and Loeb, 1995a). After transformation of the plasmid library into *E. coli* Pol $I^{ts}$ cells (Fig. 1, *step 6*) a total of 1.2 × $10^4$ cells were plated onto nutrient agar. Of these plated cells, 1400 formed colonies at 37 °C, a nonpermissive temperature for the bacterial host.

To confirm the temperature-independent phenotype of transformed cells that formed colonies at 37 °C, plasmids were prepared from 205 such colonies and retransformed individually into Pol $I^{ts}$ cells. To assess the ability of each selected HIV RT mutant to substitute for *E. coli* DNA polymerase I, the plating efficiency and colony size were assayed following an 18-h incubation at 37 °C. Of the re-transformed plasmids, 176 (86%) showed greater than 75% plating efficiency (number of colonies at 37 °C relative to the number at 30 °C). Approximately 3% of transformants formed no colonies at 37 °C; this false positive value presumably reflects the background growth of Pol $I^{ts}$ cells observed on high cell density plates (Kim and Loeb, 1995a). In separate experiments, Pol $I^{ts}$ cells transformed with pBK18 (the stuffer plasmid without HIV RT) showed plating efficiency of 1.5% at 37 °C. Of the 176 retransformed cells that showed plating efficiency greater than 75%, 57 (32%) formed large colonies at 37 °C equivalent in size to that of Pol $I^{ts}$ cells expressing wild-type HIV RT (+++), 81 (46%) formed colonies of medium size (++), and 38 (22%) produced small colonies.

*Sequence Alterations in Mutant HIV RTs*—The nucleotide and amino acid substitutions in 109 mutants selected at 37 °C, and in 18 non-selected mutants, were ascertained by DNA sequencing (Table I). Given an oligomer 36 nucleotides in length containing 12% random substitutions at each position, there should be an average of 4.3 nucleotide substitutions in the non-selected inserts; we detected an average of 5.0 substitutions per insert, suggesting that the oligomer did in fact contain approximately 12% random substitutions. Among selected mutants, the average number of nucleotide substitutions was 3.3 (Table I), suggesting that not all substitutions produced an active HIV RT that can complement the Pol $^{ts}$ phenotype. The average number of amino acid substitutions encoded by the non-selected and selected inserts was 4.1 and 2.6, respectively. Twelve of the 109 positive clones yielded inserts

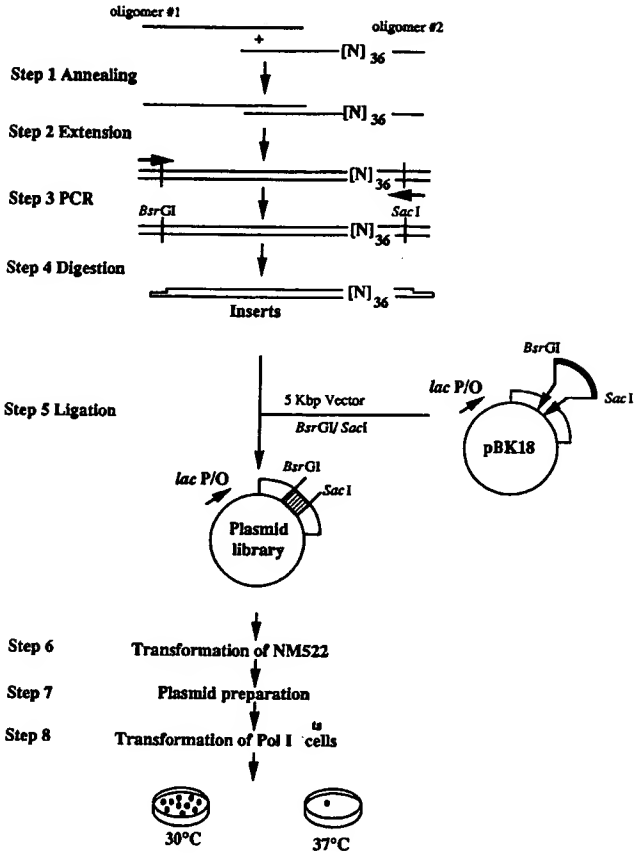



FIG. 1. **Scheme for random sequence mutagenesis of HIV RT coupled with genetic selection for functional mutants.** Construction of a plasmid-borne library of HIV RT variants containing random nucleotide substitutions at codons 67 through 78, and selection of active mutants by complementation of the temperature-sensitive phenotype of an *E. coli* DNA polymerase I mutant, is described under "Experimental Procedures." Synthesis of random nucleotide-containing oligonucleotides (*insets*) is illustrated in *steps 1–4*; $[N]_{36}$ in oligomer 2 denotes 36 contiguous residues containing 12% random nucleotides at each residue. Ligation of the inserts into the stuffer plasmid pBK18, to replace the wild-type HIV RT sequence at codons 67–78 and bring HIV RT under control of the *lac* promoter, is shown in *step 5*. Preparation of the plasmid library and selection for functional mutants in Pol $I^{ts}$ cells is shown in *steps 6–8*.

that encoded no amino acid substitutions, including 6 inserts which had nucleotide substitutions (Table I). Among 97 active mutants, some contained as many as 6 amino acid substitutions within the randomized segment.

The levels of substitution among active mutants, together with the associated complementation efficiencies, are listed in Table I. Overall, there were fewer amino acid changes among mutants that exhibited high efficiency of complementation (large colony size equivalent to that of wild-type HIV RT) than among mutants with lower complementation efficiency. Thus, the average number of amino acid substitutions in the 37 mutants that formed large colonies at 37 °C (+++) was 1.0. Among the 49 mutants that formed medium sized colonies (++), the average number of amino acid substitutions was 3.0. Among the 23 mutants forming small colonies (+) the average number of substitutions was 4.1. All of the single mutants formed colonies that were the same size as that of wild type, whereas most of the active mutants having more than 4 amino acid changes produced smaller colonies at 37 °C.

*Single Amino Acid Substitutions*—Single amino acid substitutions were observed at 9 of the 12 positions randomized

TABLE I
*Characterization of HIV RT mutant sequences*

(A) Average number of substitutions in selected and non-selected mutants; (B) distribution of amino acid substitutions in selected mutants; (C) complementation efficiency and mean number of mutations.

A

| | Nucleotide substitutions per sequence | Amino acid substitutions per sequence |
|---|---|---|
| Randomized oligomers (theoretical) | 4.3[a] | |
| Non-selected mutants ($n$ = 18) | 5.0 | 4.1 |
| Selected mutants ($n$ = 109) | 3.3 | 2.6 |

B

| Number of substitutions | Number of mutants |
|---|---|
| 0 | 12 |
| 1 | 18 |
| 2 | 23 |
| 3 | 24 |
| 4 | 21 |
| 5 | 9 |
| 6 | 2 |

C

| Complementation efficiency[b] | Nucleotide substitutions per mutant | Amino acid substitutions per mutant |
|---|---|---|
| +++ ($n$ = 37) | 1.3 | 1.0 |
| ++ ($n$ = 49) | 3.9 | 3.0 |
| + ($n$ = 23) | 4.6 | 4.1 |

[a] 12% random substitutions at 36 positions = 0.12 × 36 = 4.3.
[b] +++, large colonies equivalent to wild type; ++, medium colonies; +, small colonies.

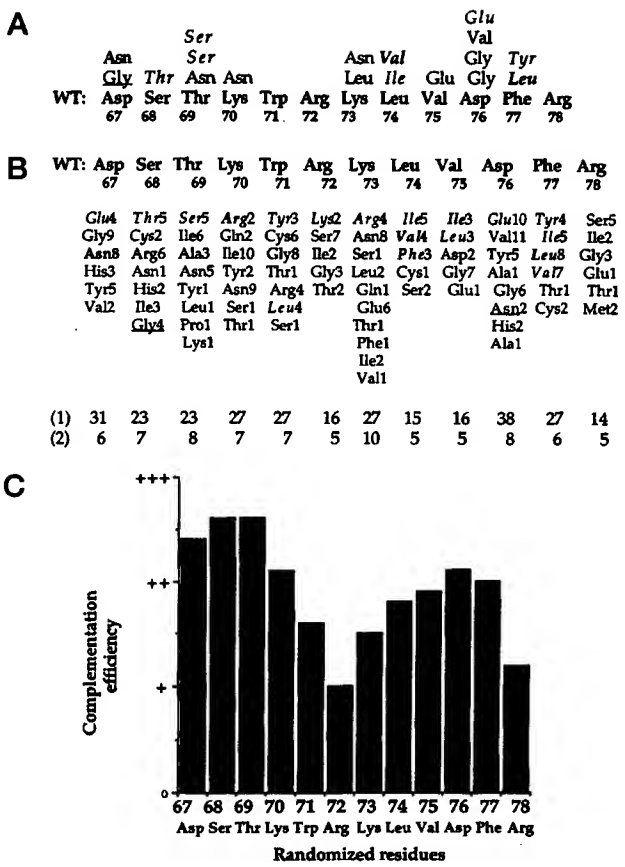



FIG. 2. **Amino acid substitutions in functional HIV RT mutants.** The amino acid sequences of 109 mutants that complemented the temperature-sensitive phenotype of Pol I[ts] cells were determined by DNA sequencing of the random nucleotide containing inserts. *Italicized* type indicates substitution of an amino acid similar to the wild-type. *Underlining* indicates a naturally occurring variation. **Bold** type indicates a substitution found, either singly or together with others, in drug resistant mutants. In *B*, the *number* to the *right* of individual mutant amino acids indicates the number of occurrences at the designated position. The number given at (*1*) below each position is the total number of substitutions at that position; the number at (*2*) is the number of different amino acids identified at that position. In *C*, complementation efficiency denotes the average colony size relative to wild type of all mutants containing substitutions at the designated residue; wild-type efficiency is denoted by +++ on the vertical scale.

(Fig. 2*A*). Mutations at 8 of these 12 positions have been observed in natural variants (positions 68 and 76) or in drug-resistant mutants (positions 67, 69, 70, 74, 75, and 77). As yet, mutations have not been found among natural variants or drug-resistant mutants at Trp-71, Arg-72, Lys-73, and Arg-78. Interestingly, Arg-71, Lys-72, and Trp-78 yielded no single mutants in this study.

*Multiple Amino Acid Substitutions*—Although individual randomized codons displayed varying yields of single amino acid substitutions (Fig. 2*A*), all 12 codons exhibited nearly equal probability of inclusion among multiple substitutions that complemented the Pol I[ts] phenotype (Fig. 2*B*). Most of the variant amino acids (italicized) have properties similar to those of the wild-type. Three mutations that were previously identified in drug-resistant mutants, D67N (AZT), K70R (AZT), and L74V (ddI), were isolated as either single or multiple mutations. In addition, we isolated two mutations, V75I and F76L, that were present as one of multiple substitutions in HIV RT from patients treated with multiple nucleoside inhibitors (Shirasaka *et al.*, 1995). In Fig. 2*C*, we show the average complementation efficiency (colony size) of all mutants that have substitutions at each of the randomized positions. Notably, mutations at positions that failed to produce single amino acid substitutions (71, 72, and 78) were associated with reduced complementation efficiency at 37 °C, suggesting that these residues are important for DNA-dependent DNA polymerase activity.

*Purification of HIV RT Mutants*—Wild-type and 13 mutant HIV RTs that complemented the Pol I[ts] phenotype were purified to near homogeneity. The mutants, all of which exhibited >75% plating efficiency and contained from one to six amino acid substitutions, are listed in Table II, together with complementation efficiency (colony size). The HIV RTs, expressed in *E. coli* as $NH_2$-terminal fusions with a "histidine tag," were purified by metal chelation chromatography. Approximately 95% of the purified wild-type preparation consisted of a 66-kDa protein (p66 subunit) and 5% of a 51-kDa protein (p51 subunit), estimated by visual inspection of a Coomassie Blue-stained 12% SDS-polyacrylamide gel (data not shown). Ten of the 13 mutant HIV RT preparations were similar in yield and purity to the wild-type (0.25–1 mg of total protein of 90–95% purity). Three of the mutant preparations were less pure (about 0.1 mg of total protein of 50–75% purity); reduced recovery of these three proteins could be due to instability, possibly correlated with an increased number of amino acid substitutions (5 to 6 per mutant). Presumably, the restricted time of induction with IPTG (2 h) limited proteolytic processing of p66 to p51.

*DNA Polymerase Activity of Mutant HIV RTs Mutants*—DNA-dependent DNA polymerase activity of purified HIV RTs, assayed on a gapped DNA template, is listed in Table II. These data demonstrate a positive correlation between complementation efficiency, estimated as colony size, and *in vitro* DNA polymerase activity. Thus the six single and two double mutants displayed a wild-type level of complementation efficiency together with enzymatic activity between 40 and 100% of wild-

TABLE II
*DNA polymerase activity of purified HIV RT mutants*

| Mutant | Complementation efficiency[a] | Activity[b] |
|---|---|---|
| | | % |
| WT | +++ | 100 |
| D67G | +++ | 90 |
| **D67N** | +++ | 85 |
| T69N | +++ | 90 |
| L74I | +++ | 100 |
| F77Y | +++ | 100 |
| F77V | +++ | 75 |
| **L74V**/V75L | +++ | 40 |
| **D67N**/K73T | ++ ~ +++ | 45 |
| **D67N**/W71G/K73N | + ~ ++ | 15 |
| **D67N**/R72S/K73E | ++ | 10 |
| R72T/L74V/D76Y/R78T | + | 5–10 |
| D67H/S68R/K70R/W71S/L74F | + | 5 |
| D67A/W71L/K73H/L74I/D76A/F77I | + | 15 |
| D67G/S68T/T69S/R72K/V75I | + ~ ++ | 25 |

[a] +++, large colonies equivalent to wild type; ++, medium colonies; +, small colonies.
[b] DNA-dependent DNA polymerase activity was assayed as incorporation of label from [*methyl*]-$^3$H]TTP into activated DNA by 0.1 μg of purified protein at 37 °C for 20 min. Activity was normalized to that of wild-type HIV RT.

type. The 6 remaining mutants with larger numbers of substitutions displayed decreased complementation efficiency together with 5–25% of wild-type activity (Table II).

*Inhibition of Mutant HIV RTs by AZTTP*—AZT sensitivity, measured as inhibition of DNA-dependent DNA polymerase activity by AZT triphosphate, was determined for eight highly active mutants (D67N, D67G, T69N, L74I, F77Y, F77V, LV74/V75L, and D67N/K73T). Among the single mutants, a 18- and 16-fold decrease in $IC_{50}$ for AZT triphosphate was observed for F77Y and F77V, respectively, relative to wild type (Fig. 3*A*). L74I showed a 13-fold decrease in $IC_{50}$ (Fig. 3*B*). D67G showed an approximately 4-fold decrease, while no reduction in sensitivity was observed with two other substitutions at the same position (D67N and T69D). Previous studies with purified HIV RT from AZT-resistant mutants (*i.e.* D67N, K70R, T215Y, and K219Q) failed to demonstrate *in vitro* drug resistance even though viruses containing the same substitutions together with others showed a large increase in the $IC_{50}$ for AZT in cultured cells (Larder and Kemp, 1989). Consistent with these observations, D67N/K73T showed decreased sensitivity to AZTTP even though D67N alone did not.

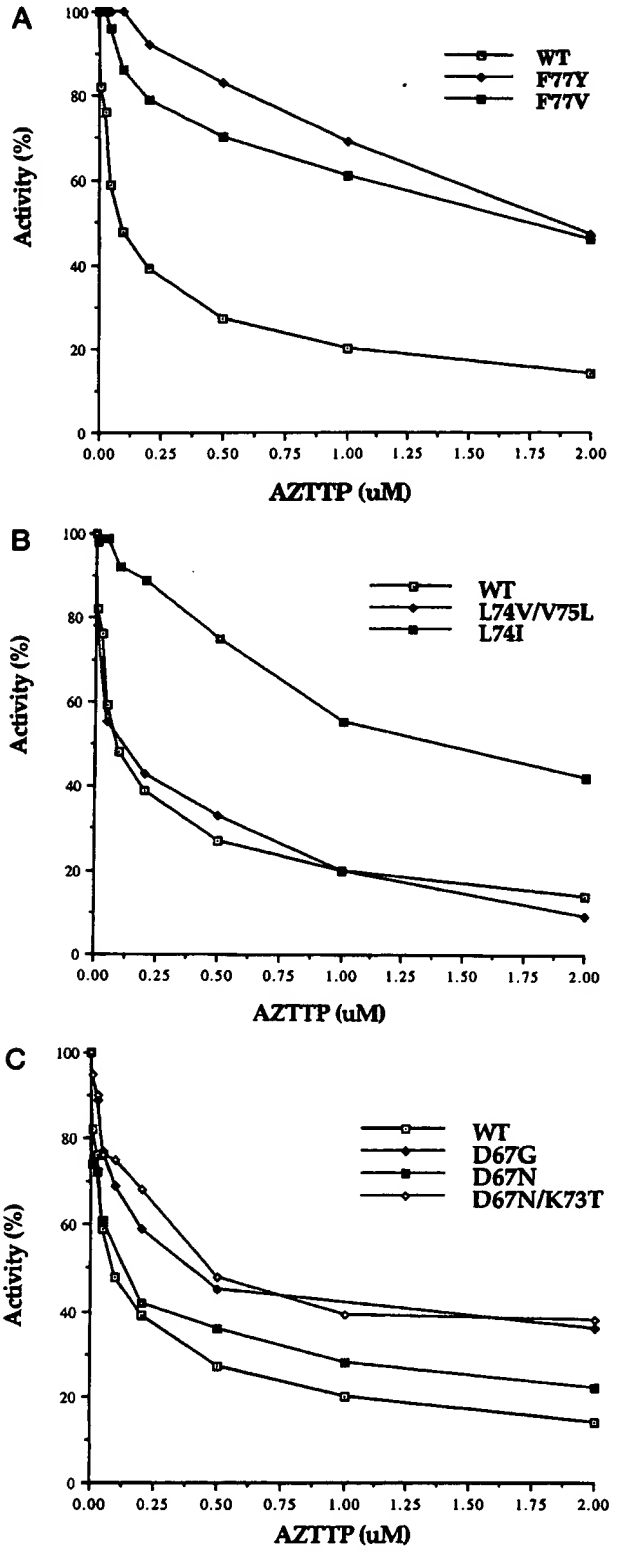



FIG. 3. **Inhibition of DNA polymerase activity of HIV RT mutants by AZTTP.** The DNA-dependent DNA polymerase activity of HIV RT mutants was assayed on a gapped DNA template in the presence of AZTTP as described under "Experimental Procedures."

## DISCUSSION

In this work, we tested whether complementation of a replication defective *E. coli* DNA polymerase I mutant by HIV RT (Kim and Loeb, 1995a) can be used to select active HIV RT mutants from a population of variants containing random nucleotide substitutions (Horwitz and Loeb, 1986). In our functional complementation system, expression of HIV RT enables an *E. coli* mutant harboring temperature-sensitive DNA polymerase I to grow at a nonpermissive temperature (Kim and Loeb, 1995a). In these initial experiments with random sequence substitutions in HIV RT, we targeted a region that has little secondary structure and would likely permit a large number of substitutions that yield functional mutants. The crystal structure of HIV RT indicates that the 12 amino acid target forms a flexible loop between the β3 and β4 domains (Tantillo *et al.*, 1994). In addition, many viral isolates that are resistant to nucleoside analogs contain mutations within this region. The inherent flexibility of the target and the occurrence of variants in the natural host is consistent with our finding that all 12 residues are tolerant of substitutions, either singly or in combination with others.

We observed in the experiments described here that approximately 12% of Pol $I^{ts}$ cells transformed with a library encoding random substitutions at amino acids 67–78 formed colonies at 37 °C. The fraction of a library that supports colony formation at nonpermissive temperature, *i.e.* the fraction of encoded sequences that can complement the Pol $I^{ts}$ phenotype, depends on several factors. These include the upper and lower limits of intracellular HIV RT activity compatible with growth of the bacterial host (the window for selection), the proportion of random nucleotides relative to wild-type nucleotides within the substituted segment, and the function of the targeted amino acids in supporting DNA-dependent DNA polymerase activity. With respect to the first factor, it appears that the fraction of transformants capable of complementation can be adjusted by controlling the level of HIV RT expression. Thus, we observed that the fraction of transformants that form colonies at 37 °C decreased from 12 to 7% when IPTG was deleted from the selection plates (data not shown). By reducing the level of expression, we presumably limited recovery to mutants with relatively high specific activity. In the experiments described here where use of 1 mM IPTG gave 12% positive transformants, one-third (*i.e.* 4%) yielded colonies of wild-type size, suggestive of an intracellular HIV RT activity comparable to wild-type. That higher levels of HIV RT expression can be lethal to the host (Kim and Loeb, 1995a) indicates that there is an upper limit of activity compatible with growth. A second factor affecting the fraction of positive transformants is the degree of randomness in the targeted region, with greater randomness and more amino acid substitutions mitigating against levels of activity that promote growth and colony formation. In the present experiments, the target sequence was replaced with 12% random nucleotides at each position, and as a result the wild-type sequence should have occurred at a frequency of 1%, corresponding to 120 of the 12,000 cells plated. Assuming that the predicted 120 wild-type sequences were included in the 1400 positive transformants we recovered, we should have found 9 wild-type isolates (120/1400) among the 109 that we sequenced. That we found 6 wild-type sequences among the 109, approximately the expected occurrence, substantiates the designated composition of the library. Third, it is likely that the importance of the targeted amino acids for DNA polymerase activity affects the fraction of transformants selected, with substitution of residues having crucial functions mitigating against ability to complement. The high frequency of positive transformants that we recovered by targeting the β3-β4 flexible loop is consistent with the nonessential nature of the individual amino acids.

The amino acid replacement we observed among mutants that complement the Pol $I^{ts}$ phenotype (Fig. 2) indicates that Trp-71, Arg-72, and Arg-78 are relatively intolerant of substitution compared with other residues in our 12-amino acid target. This pattern is consistent with substitutions recovered in HIV isolates and with sequence conservation among various viral RTs. Within the target region, substitution at two positions, S68G and D76N, occurs in natural HIV RT variants, and substitution at 6 positions (Asp-67, Thr-69, Lys-70, Leu-74, Val-75, and Phe-77) occurs in drug-resistant viral isolates. Of the 4 residues (Trp-71, Arg-72, Lys-73, and Arg-78) not included among viral mutations, Trp-71, Arg-72, and Arg-78 yielded no substitutions in the single mutants we sequenced (Fig. 2*A*); moreover, mutants with multiple substitutions including one or more of these three positions showed low complementation efficiency (Fig. 2*C*). These data indicate that Trp-71, Arg-72, and Arg-78 are relatively rarely mutable and suggest that they may be important in DNA-dependent DNA polymerase activity. In fact, a recent study showed that two mutations at Arg-72, including R72K, result in greatly decreased activity due to reduction in translocation (Sarafinos *et al.*, 1995). Notably, sequence comparison among viral RTs reveals that Trp-71, Arg-72, and Arg-78, as well as Leu-74 and Asp-76 are highly conserved (Barber *et al.*, 1990). Taken together, available data indicate that mutations at Arg-72 may be incompatible with viral replication in the natural host.

Among the mutants we analyzed, the number of amino acid substitutions is inversely correlated with colony size and with DNA polymerase activity on a gapped DNA template (Table II). These observations suggest that ability to complement in the *E. coli* system reflects the DNA-dependent DNA polymerase activity of HIV RT, with mutants having 10% of wild-type DNA-dependent DNA polymerase activity being selectable. It has been reported that an HIV RT mutant with 37% of wild-type activity, L74M, supports production of viable virions in cultured human cells, whereas another mutant, L74A, exhibiting 30% of wild-type activity, does not (Lacey and Larder, 1994). Thus, selection for mutants that support bacterial growth is apparently less stringent than selection for virion production in cell culture systems. In accord, the HIV RT mutants we selected that have significantly reduced DNA polymerase activity (Table II) are unlikely to be identified in viral populations because they presumably cannot support viral replication.

Six amino acids within the segment we targeted (Arg-67, Thr-69, Lys-70, Leu-74, Val-75, and Phe-77) have been found to be mutated in variants resistant to nucleoside analogs (Tantillo *et al.*, 1994). We analyzed some of the mutations we isolated at these positions for AZTTP resistance. As illustrated in Fig. 3*A*, F77Y and F77V conferred an 18- and 16-fold decrease in inhibition by AZTTP, respectively. Although these substitutions have not been reported in viral isolates from patients, it was recently reported that individuals receiving therapy with a combination of nucleoside analogs developed a F77L mutation subsequent to a Q151M mutation. Thus, F77L is encompassed in the evolution of a set of five mutations, beginning with Q151M, followed by F77L and F116Y and later by A62V and V75I, that confers multiple drug resistance (Shirasaka *et al.*, 1995). The L74I mutation we isolated exhibited an 8-fold reduction in inhibition by AZTTP; Fig. 3*B*). A similar mutation, L74V, has been isolated from patients after ddI treatment (St. Clair *et al.*, 1991) and exhibited reduced sensitivity to ddGTP (Lacey *et al.*, 1992). Thus Leu-74 is likely to be involved directly or indirectly in substrate selection. Interestingly, L74V displayed wild-type sensitivity to AZTTP in combination with V75L. The mutation D67N, in combination with additional mutations (K70R, T215Y, and K219Q) exhibits a more than 100-fold increased resistance to AZT in a cell culture system (Larder and Kemp, 1989). However, a purified HIV RT mutant containing all four mutations failed to show resistance to AZTTP *in vitro* (Skalka and Goff, 1993). In our study, D67N did not affect sensitivity to AZTTP, consistent with the observation that virions containing D67N as a single mutation did not show increased resistance to AZT *in vivo* (Larder *et al.*, 1991; Larder and Kemp, 1989). However, we did find that both D67G and D67N/K73T exhibited increased resistance to AZTTP (Fig. 3*C*).

Many mutations in HIV RT have been observed in virions that are resistant to nucleoside analogs in patients and in cultured cells. However, purified HIV RT bearing the same mutations frequently does not exhibit resistance to the corresponding nucleoside triphosphates when assayed *in vitro*. Conversely, we isolated both single and multiple substitutions that render HIV RT resistant to AZTTP, but have not yet been identified in drug-resistant viral isolates. Considering the exceptionally high mutation rate of the virus (Preston *et al.*, 1988; Roberts *et al.*, 1989), the high error rate of the reverse tran-

scriptase *in vitro* (Preston *et al.*, 1988) and recent evidence for rapid viral replication during the course of HIV infection (Delwart *et al.*, 1993), it seems likely that all single base substitutions and many multiple substitutions would have arisen and been "tested" for resistance during AZT therapy in patients. If viruses containing the mutations we identified are in fact resistant to AZT, why have they not been selected? Some may not have a high enough DNA-dependent DNA polymerase activity to compete successfully (*e.g.* F77V, Table II). However, some (*e.g.* F77Y and L74I which may have wild-type levels of this activity) might interfere with additional activities or properties of HIV RT which are essential for production of virions, such as strand transfer/displacement, processivity, or replication fidelity, and might be inadequate for viral replication in the natural host.

Several conclusions can be drawn from our initial experiments on random sequence substitution of HIV RT and selection by genetic complementation. First, our positive genetic selection system can identify active HIV RT mutants. The mutants we characterized exhibit from 5 to 100% of the DNA-dependent DNA polymerase activity of the wild-type enzyme and include a majority of substitutions located within the target region that have been observed as natural variants and drug-resistant mutations. The ability to select active mutations and to characterize their phenotypes in an alternate bacterial host provides a powerful means for assessing the consequences of specific amino acid substitutions, and combinations of substitutions, on DNA polymerase activity. Second, random mutagenesis of HIV RT enables us to assess the mutability of each amino acid residue in the enzyme. Rarely mutable or immutable residues can be identified by this approach, and such residues can serve as potential targets for more efficacious anti-HIV drug therapy that precludes or delays emergence of resistant variants. Since random mutagenesis can be applied to large targets in HIV RT, and can also survey all possible combinations of amino acid substitutions within the target, the approach offers different parameters for identifying essential amino acids than does inactivation by substitution with alanine (Richardson and Richardson, 1990). Although substitutions within the flexible region we targeted in this study did not reveal immutable sites, it remains to be determined if other more structured domains contain such immutable (essential) residues. Finally, random mutagenesis creates HIV RT mutants which are unlikely to be recovered from the natural host because they cannot support viral replication. Phenotypic and biochemical analysis of such mutants, many of which will have new or altered biochemical properties, can provide insight into the involvement of individual amino acid residues in catalysis and into structure-function relationships within the enzyme.

*Acknowledgments*—We are grateful to Ann Blank and Margaret Black for generous counsel throughout the course of this study.

REFERENCES

Battula, N., and Loeb, L. A. (1974) *J. Biol. Chem.* **249,** 4086–4093
Boucher, C. B., O'Sullivan, E., Mulder, J. W., Ramantarsing, C., Kellam, P., Darby, G., Lange, J. M. A., Goudsmit, J., and Larder, B. A. (1992) *J. Infect. Dis.* **165,** 105–110
Boyer, P. L., Tantillo, C., Jacobo-Molina, A., Nanni, R. G., Ding, J., Arnold, E., and Hughes, S. H. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 4882–4886
Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
Delwart, E. D., Shpaer, E. G., Louwagie, J., McCutchan, F. E., Grez, M., Rubsamen-Waigmann, H., and Mullins, J. I. (1993) *Science* **262,** 1257–1261
Fitzgibbon, J. E., Howell, R. M., Haberzettl, C. A., Sperber, S. J., Gocke, D. J., and Dubin, D. T. (1992) *Antimicrob. Agents Chemother.* **36,** 153–157
Gu, Z., Gao, Q., Li, X., Parniak, M. A., and Wainberg, M. A. (1992) *J. Virol.* **66,** 7128–7135
Gu, Z., Fletcher, R. S., Arts, E. J., Wainberg, M. A., and Parniak, M. A. (1994) *J. Biol. Chem.* **269,** 28118–28122
Horwitz, M. S., and Loeb, L. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 7405–7409
Kim, B., and Loeb, L. A. (1995a) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 684–688
Kim, B., and Loeb, L. A. (1995b) *J. Virol.* **69,** 6563–6566
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 488
Lacey, S. F., and Larder, B. A. (1994) *J. Virol.* **68,** 3421–3424
Lacey, S. F., Reardon, J. E., Furfine, E. S., Kunkel, T. A., Bebenek, K., Eckert, K. A., Kemp, S. D., and Larder, B. A. (1992) *J. Biol. Chem.* **267,** 15789–15794
Larder, B. A., and Kemp, S. D. (1989) *Science* **246,** 1155–1158
Larder, B. A., Kellam, P., and Kemp, S. D. (1991) *AIDS* **5,** 137–144
Mitsuya, H., and Broder, S. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 1911–1915
Preston, B. D., Poiesz, B. J., and Loeb, L. A. (1988) *Science* **242,** 1168–1171
Richardson, J. S., and Richardson, D. C. (1990) *Principles and Patterns of Protein Conformation,* Plenum Press, New York
Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A., and Kunkel, T. A. (1989) *Mol. Cell. Biol.* **9,** 469–476
Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Reference Manual,* Cold Spring Harbor Press, Cold Springs Harbor, NY
Sarafianos, S. G., Pandey, V. N., Kaushik, N., and Modak, M. J. (1995) *J. Biol. Chem.* **270,** 19729–19735
Shirasaka, T., Kavlick, M. F., Ueno, T., Gao, W.-Y., Kojima, E., Alcaide, M. L., Chokekijchai, S., Roy, B. M., Arnold, E., Yarchoan, R., and Mitsuya, H. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 2398–2402
Skalka, A. M., and Goff, S. P. (1993) *Reverse Transcriptase,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
St. Clair, M. H., Martin, J. L., Tudor-Williams, G., Bach, M. C., Vavro, C. L., King, D. M., Kellam, P., Kemp, S. D., and Larder, B. A. (1991) *Science* **253,** 1557–1559
Tantillo, C., Ding, J., Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H., Pauwels, R., Andries, K., Janssen, P. A. J., and Arnold, E. (1994) *J. Mol. Biol.* **243,** 369–387
Witkin, E. M., and Roegner, M. V. (1992) *J. Bacteriol.* **174,** 4166–4168